

Bush badly needs good advice

President George Bush has promised to create better machinery for gathering technical advice, but time is passing. A search for a paragon should not let him hide from the urgency of his need.

THE new president of the United States may be saying all the right things (see *Nature* 337, 585; 1989), but he is doing fewer of them than he should. This complaint has been levelled generally at the new administration, which has yet to fill its roster of cabinet appointments, let alone all the second-rank posts, all of whose incumbents will have to be interviewed by a committee of the Senate and judged capable of doing their jobs. But the complaint is particularly apt when applied to President Bush's plans for the more effective administration of science, among the most ambitious of those put forward last month.

Briefly, Bush would give a new science adviser the status of "assistant to the President", putting him or her on a par with the head of the National Security Agency (now General Brentt Scowcroft). On top of that, there is to be a Council of Science and Technology Advisors resembling in all but name what used to be the President's Science Advisory Committee (PSAC), much respected until killed off by President Richard M. Nixon. But nobody has yet been nominated for the top job. Even if a name should have appeared before the end of this week, a quarter of a year (or a sixteenth of a presidential term) will have gone before the man or woman can set to work.

The administration's general excuse for its tardiness — that its succession of the like-minded Reagan administration means fewer policy discontinuities and thus a lesser need of new officials — does not apply to the science advisory job which, as advertised, is an innovation. The present holder of the office, Mr William Graham, has had only meagre influence during his two years, and is unlikely to be struggling for promotion.

So why not pick somebody and get on with it? The word is that the administration is looking for a science adviser from industry, which is consistent with the Bush view of science as a means of improving the competitiveness of the United States in the "global marketplace". One snag is the formidable list of loyalty tests the new incumbent will have to survive: finding a Republican will be easy, finding somebody willing to go to the stake for the Strategic Defense Initiative will be something else.

A more practical difficulty may be that there is already a technical person powerfully placed at the White House. Mr John Sununu, Bush's chief of staff, was a chemical engineer before becoming governor of New Hampshire. Between gaps in what must be the fullest appointments diary in the Western world, not to mention side-trips to

Japan, China and Korea, Sununu is interviewing potential candidates for the new job. Quite apart from his lack of time, there may be a further difficulty: some putative incumbents may be deterred by knowing that the man who allocates space in Bush's working day is likely to have his own strong views on what science advice the president should be given.

Yet the need for an appointment becomes more urgent every day. In just over two months, the National Security Council is to have completed a review of US policy on national security, to which an able science adviser would have a powerful contribution to make.

Soon afterwards, early in the summer, there will be a ministerial meeting of the North Atlantic Treaty Organization that will be crucial for the future of US weapons policy. Meanwhile, the administration is being challenged by the daunting character of its own agenda (improving public education, doing something effective about AIDS and drugs) as well as by the conundrums thrown up by the Congress, the still-unfocused worry about climatic change for example (see page 3). The danger is that too many of these issues will be decided by default.

The reconstitution of PSAC will raise other problems. From its creation after the Second World War until its abolition, the old PSAC won respect by the self-evident independence of its members. Sometimes it could be an irritating talking-shop, but its strength was its readiness to oppose administration policy when it thought fit. Bush is looking for a council of "leading scientists, engineers and distinguished executives from the private sector" (which includes the universities), which is not in itself a bad prescription. But will he allow that the advisers should be chosen for the calibre of the advice they have to offer, not the likelihood that the advice will be acceptable? That will be the acid test of whether Bush means all the right things he is saying. □

Science squeezed

Nobody should be surprised that ten years of short commons have deprived Britain of researchers.

THE emergence in Britain of a shortage of young researchers (see page 7) is about as remarkable as the diurnal rhythm. For close on ten years, British research



establishments have been harassed by lack of funds, so that newspapers have been full of tales of researchers who have been put out to grass or who have fled abroad. Young people, who read the newspapers, have also learned from them that entrants to other professions, that of providing financial services, for example, are likely to live more prosperously, and have voted with their feet. That there is a shortage of putative graduate students is not in the least surprising, but a predictable consequence of the market forces on which the British government bases most of its policies, sometimes successfully. It is even more worrying that the past few years have probably frightened away an unduly large proportion of the most able people.

There is now no quick remedy. The Medical Research Council's plan to offer graduate students on its books an extra £500 a year, commendable enough, is about as relevant to the underlying market forces as would be a great man's attempt to interfere with the diurnal tides by walking into the sea (which is what King Canute demonstrated nine centuries ago). Even doubling the present meagre stipend would have little beneficial effect. Only when the practice of research in Britain has been shown again to be an honoured as well as an honourable profession will young people be clamouring for the studentships on offer.

That calculation should be central in the planning of how to spend the extra sums for research now being made available. But that appears not to have been the case. Too much of the extra is being spent on directed research or, worse, on the research councils' own establishments, too little on the means (research grants) by which individuals might make a mark and thus fire the imagination of young people choosing a career. Researchers should fight to see this balance redressed, knowing that in doing so they will be serving the wider British national interest.

Meanwhile, as a stop-gap remedy, the research councils might think of offering their studentships to young people from abroad, coupled with work permits to stay in Britain permanently. In the United States, for different reasons afflicted by a shortage of young researchers, the stragem works well enough. □

Satanic violences

The Shiite threat to kill a British author and the violence of animal rights extremists have a lot in common.

MR Salman Rushdie's book *Satanic Verses* opens with the fantasy of how two unremarkable people survive the destruction (by a bomb) of an aircraft over the English Channel and then found a new religion. The ways of the Shiite Muslims who have now offered a money prize for Rushdie's assassination and those of the people who advocate civil rights for animals may seem only remotely connected, but they are nevertheless worth remarking.

Apart from the bomb last week at the University of

Bristol (remarkable only for being the first use of high explosives to damage British property in this particular cause) and the occupation of a crane on a Berkeley construction site, this has been a relatively quiet week, yet in their character the protests on behalf of animals have precisely the qualities that make the threats against Rushdie as offensive as they are.

First, it must be clear that the Shiite threat is an act of violence. One apologist last week is quoted as saying that "the arrow . . . is already travelling to his heart", by which is presumably meant that the threat will be regarded for decades to come as an invitation to some impressionable assassin to push the arrow home. The more tangible violence of the extremists in the animal rights movements has the same implacable quality; nothing, especially not reason, can divert it. It has the same quality of detachment from the identities of those who are injured; Shiites imprison random Western travellers to Beirut for years on end, while the explosion of a bomb at the University of Bristol is similarly random (but there is a veterinary school there). Why pick on that rather than on some other British university? Because it happened to be handy, perhaps?

As with the case of Rushdie so on the animal rights front, the most urgent need is that the moderates who believe that Rushdie's book is blasphemous (to Muslims) or that animals are misused in research should clearly dissociate themselves from the actions of the extremists. At least in the use of animals in research, the moderates have a powerful case, not least in their demands that people's use of animals should be as sparing as possible and that humane procedures should be followed. Failure to condemn these acts of violence has two consequences; the extremists are excited, while the moderate case is so tarnished by its association with violence that it is dangerously weakened. Sadly, that lesson seems not yet to have been learned, either in the case of Rushdie's book or that of the bomb at Bristol. □

Unfamiliar disguises

The unfamiliar appearance of this issue of *Nature* implies no systematic change of content.

REGULAR readers who happen to notice the typographical changes in this week's format will passionately resent the small changes which have been made. That at least is what past experience shows. The only comfort is that, with the passage of a few weeks, the same readers will become just as vigilant in the defence of what is now offered — a layout which is generally more uniform, in particular respects more workable and which may help to make some sections of *Nature* more readable. No attempt has been made, on this occasion, to introduce fancy typographical tricks or even to restore the couplet from Wordsworth's "Endymion" from which this journal's name was taken in 1869. □

US Congress plans greenhouse legislation

- Tough measures on global warming
- Scientific opinions differ

Washington

THE new US Congress has begun with a stampede by legislators introducing bills meant to respond to the perceived threat of global warming. But researchers at congressional hearings last week seemed much more cautious than the legislators in estimating what the future holds for the world's climate.

Several wide-ranging bills on global warming have been introduced since the Congress opened on 25 January. The most recent, the "Global Warming Prevention Act of 1989", was announced by Claudine Schneider (Republican, Rhode Island) of the House of Representatives on 21 February, and has 50 co-sponsors.

Schneider's bill embodies many of the elements of those launched a few weeks ago by Senator Timothy Wirth (Democrat, Colorado) and Senator Albert Gore (Democrat, Tennessee). It calls for an international agreement on a 20 per cent reduction in global carbon dioxide concentrations by the year 2000 and the revision of the Montreal Protocol to phase out all chlorofluorocarbon production within five to seven years. Under the legislation, US aid for family planning services in developing nations is doubled to help reduce population growth (and hence energy demand) and recipients of foreign aid will be required to practise sustainable forest management. Tropical timber imports will be banned from countries that do not comply.

On the domestic front, the bill would provide hundreds of millions of extra dollars for research and development on energy-efficient technology, solar and renewable energy resources and solar-generated hydrogen fuels. There would be tax rebates for fuel-efficient cars, and increased 'gas guzzler' taxes.

The drive to legislate stems from widespread public concern about global warming, and the greenhouse effect. The issue caught the attention of the media during last summer's drought in the United States. But, ironically, few if any researchers attribute the drought to the greenhouse effect. Stephen Schneider of the National Center for Atmospheric Research (NCAR), giving evidence to the House of Representatives' subcommittee on Energy and Power last week, said it is "absurd" to suggest a direct causal link. But he welcomes the attention global warming is receiving in Congress even if it is "for the wrong reasons".

NCAR's Schneider strongly suspects

that the high average global temperatures of the 1980s, which are about half a degree Celsius higher than at the end of the last century, are a greenhouse signal; he advocates action now instead of later.

But Robert Correll, of the National Science Foundation, in evidence last week to the Senate Commerce, Science and Transportation Committee, noted that the increase of temperature over the past century has not been steady, in contrast to model predictions of the greenhouse effect. And Correll says that the current differences of opinion among scientists as to whether there is now a greenhouse

warming are "substantial".

Some members of both committees seemed irritated by the lack of consensus and warned the scientific witnesses that their uncertainty might be used by some congressmen as an excuse for no action.

Regardless of uncertainty, opposition to the new legislation is bound to be strong. The powerful automobile industry will oppose legislation that favours more fuel-efficient imports, and the coal industry will object to action that might substantially reduce coal consumption. The new bills are expected to have a particularly rough passage through the Senate Energy Committee, where these interest groups have a strong voice. But there is considerable public pressure on Congress to "do something", and if the heat waves strike again this summer — particularly if they strike Washington — it is almost certain that some of the new legislation will be passed.

David Swinbanks

■ See also pages 15 and 54.

US REACTORS

Shoreham plant to be revived?

Boston

NEW YORK'S never-licensed Shoreham Nuclear Power Station may be rising from its grave, if its parent utility's stock is any indication. The stock of the Long Island Lighting Company (LILCO) has recently been the third most actively traded of the stocks, on the New York Stock Exchange.

The speculative flurry has been caused by two developments in the tangled but continuing saga of the nuclear plant. A federal judge has thrown out a case brought by the local legislature which accused LILCO of "racketeering" and corruption, and the US Nuclear Regulatory

also have raised electricity rates, even though it would have restricted the utility's right to pass on costs to ratepayers (see *Nature* 333, 588; 1988).

The plan was sanctioned by all the major parties involved, but was derailed unexpectedly when the New York State legislature refused to ratify it out of fear of political reprisals.

Ironically, Long Island constituents, whose electricity rates are already the second highest in the United States, will now face even greater cost increases, whether or not the plant goes into operation.

There are still many hurdles ahead for the Shoreham reactor, most notably the strong local and state-based political opposition.

Shoreham's local Suffolk County legislature voted unanimously to appeal against the rejection of its racketeering and corruption suit against LILCO, turning down a settlement of nearly \$400 million offered by the utility in reduced electric rates. The legislature said it would not settle the case unless the utility promised that it would "irrevocably close, shutdown and decommission" Shoreham.

But despite their differences, all the parties involved agree that Shoreham's fate is highly uncertain. Even LILCO representatives say that they are now following a "two-track" strategy: moving ahead to license the reactor, and trying to negotiate a scheme to shut the plant down. What they find intolerable is the waiting and indecision, which they claim has cost the utility \$1.700 million in interest and finance charges since the plant was completed.

Seth Shulman



Commission is expected to rule soon upon whether to grant the Shoreham reactor a low-power operating licence. Both developments promise a future for the power station, which was completed four years ago.

The Shoreham plant was widely pronounced dead last spring, when New York State authorized a plan to close and decommission the reactor, but which would

University building bombed

London

BRITISH police now fear a bombing campaign against other universities after the explosion of a bomb, thought to have contained 5 pounds of high explosives, in an administrative building of the University of Bristol last week. The bomb went off in the staff restaurant of the Senate House at midnight on Wednesday, 22 February. There were no casualties.

Responsibility for the attack has been claimed by the Animal Liberation Front (ALF), a well-known group of extremists, and by two previously unknown groups, the Animal Defence Organisation and the Animal Abused Society, which warned the police beforehand of the bomb. The ALF has usually concentrated its attacks on the houses of scientists, which have been daubed with paint, and on retail shops selling items such as furs.

The Secretary of State for Education and Science, Mr Kenneth Baker, visiting the university after the explosion, called the attack "an appalling example of terrorism". He said that tighter security might be necessary at British universities.

British animal rights groups have previously had some success with incendiary

Bomb damage inside the Senate House.

devices, but have not hitherto used high explosives. Animal rights groups which campaign using non-violent means against vivisection have condemned the bombing. A spokesman for the British Union for the Abolition of Vivisection says it will cause extreme damage to the campaign against animal experimentation. There are 50,000 members of non-violent groups, he says, and less than 100 in the loosely-organized group of individuals which makes up the ALF.

Christine McGourty

Activists setting the agenda

Berkeley

THE divisive issue of the use of animals in research was the centre of attention last week at two San Francisco area universities. Stanford University has been debating a genteel change of policy that would affect the use of animals in teaching, while, on the University of California campus at Berkeley, a more acerbic debate over an animal care facility has flared up.

Animal rights activists have been stepping up their activities in California. Last week, six of them scaled a 175-foot crane at a construction site on the Berkeley campus to protest at the building of the

university's new animal facility. Unfurling banners that read "Stop Germ Warfare Lab!" and "No Toxic Animal Lab!", the protesters vowed to remain on the crane for two weeks. But university officials have called the stunt "ironic", because the crane is not involved in the construction of the new animal facility, but instead is being used to put the finishing touches to a plant-biology building.

The university is hesitant to remove the protesters, fearing that somebody would be hurt in the process. Construction of the animal facility has continued undisturbed, but the protest is costing the university more than \$10,000 a day, according to

one spokesman, to pay for 24-hour security guards and lost construction time.

Meanwhile, in an effort to avert future animal rights controversies in the classroom, Stanford University has drafted a policy that would inform students which courses and degree programmes use animals in instruction. The new policy, now being finally revised, encourages students to raise with their instruc-

David Maung/Daily Californian

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

High-level protests against Berkeley's new laboratory.

UCT again confronts government

Oxford

THE University of Cape Town (UCT) is once again in confrontation with the South African government, this time because of a breach of confidence by the Minister of Health and Population Development, Dr Willie van Nierkerk.

The issue concerns Dr Jocelyn Kane-Berman, whom van Nierkerk sacked from her former post as superintendent of the Groote Schuur Hospital, UCT's teaching hospital, and the post of dean of the medical school at UCT, which the university is now seeking to fill against a pending retirement. In reply to parliamentary questions about the dismissal, van Nierkerk said that he would approve Kane-Berman's appointment as dean because, at UCT, she would be "in an environment where the proximity of the ANC (African National Congress) is well-known".

Apart from the slur on the university, UCT is incensed that van Nierkerk should comment publicly on one application for an academic post which is still under consideration. Van Nierkerk, a former professor of obstetrics and gynaecology at Stellenbosch, can hardly be ignorant of the niceties of academic appointments.

UCT now says that the insinuations in the minister's remarks are "disgraceful". Dr Stuart Saunders, the vice-chancellor, and Mr Leo Abrahamse, the chairman of the council, who met the minister to protest, said that van Nierkerk had chosen not to provide an explanation, and that the university would now be making representations through other channels.

The cause of Kane-Berman's sacking from Groote Schuur was her statement, to a weekend newspaper, that Nelson Mandela would be her choice as prime minister in a South African government chosen on merit (see *Nature* 336, 612; 1988). The *South African Medical Journal* has called for her reinstatement, but the minister is unrepentant. Meanwhile, Kane-Berman intends to sue her former employers, the provincial administration, for reinstatement.

Michael Cherry

tors any concerns they may have about animal use. But although it provides room for individual solutions, the policy does not oblige instructors to change course requirements.

Stanford spokesman Robert Beyers said the concept of forewarning students was stimulated by the passage in the state legislature last spring of a law that allows primary and secondary school students to refuse to dissect animals in the classroom.

The new rule may have little impact, because animals are rarely used in undergraduate courses, and no required courses in the medical school or the biology department involve animals. **Marcia Barinaga**

Four groups join battle

Washington

THE US Patent Office is now wrestling with competing claims to patents on high-temperature superconductivity from four different research groups — at IBM, AT&T, the University of Houston and the US Navy. The patent office says there is a conflict over priority in the four claims to have discovered useable versions of the copper oxide superconductors. If the patents office can resolve the conflict, its decision could be a boon for the organization whose claim succeeds.

The conflict, legally called an 'interference' proceeding, has arisen because the patent applications by the four are similar, overlapping, and were filed at almost the same time. The parties to the dispute have each filed a string of patent applications for both the lanthanum copper oxide materials discovered in 1986 by Nobel prizewinners Georg Bednorz and Alex Müller, of IBM Zurich, and the yttrium-based copper oxides discovered by Paul Chu of the University of Houston and Maw-Kuen Wu of Alabama University in January 1987. It is widely believed that the yttrium compounds will have more general commercial application than the lanthanum materials.

The US Patent Office will have to decide not only whether patent applications for the lanthanum oxide superconductors can be extended to cover the yttrium compounds, but also whether applications based on mixed-phase materials of unknown crystal structure are valid (as in the case of Chu's first patent application for yttrium copper oxide material).

Information about patent applications for the superconductor materials began to come to light last summer, in Europe and Japan, where by law applications are made public 18 months after submission (or after the earliest priority date claimed). US patents are not published until granted, but several of the European and Japanese patent applications include priority claims for US patents and so disclose US filing dates.

The application from the US Naval Research Laboratory has not previously been known of, having been filed only in the United States. At the applications stage, even cases of interference are not publicized by the US patent office, but Cynthia Stevens, an IBM spokeswoman on patents, confirms the identity of the four competing applicants.

Japanese patent applications are notable by their absence in the dispute, in contrast with the common fear in the United States that Japan will be the first to exploit the commercial applications of the high-temperature superconductors. One explanation may be that the early patent applications by University of Tokyo

researchers are narrowly drawn and cover only the lanthanum oxides (see *Nature* 335, 389; 29 September 1988).

Japan's Ministry of International Trade and Industry (MITI), which has filed broad applications worldwide covering the yttrium oxides, may claim interference at a later stage, according to Christopher Vear of the British Technology Group who is monitoring patent applications for the new materials. But MITI filed after the US groups in January 1987 and does not seem to have a strong position.

Japan's real strength lies in patent applications for devices made from the new materials and processes for making the superconductors into wires and thin films. Japanese companies, in particular Sumitomo Electric Corporation, have filed hundreds of such applications in Japan and many have been combined into broad patent applications in Europe, where Japanese interests have filed most applications. According to Vear's latest figures, Japan has filed 80 applications in Europe compared with 28 for the United States, 5 for West Germany, 4 for Britain, 3 for France and 1 each for the Soviet Union and Hungary.

The only patent so far granted has been awarded by the US Patent Office to John B. Vander Sande and Gregory J. Yurek at the Massachusetts Institute of Technology (MIT) for a process that combines the new ceramic materials with silver to make a more malleable material (see *Nature* 336, 607; 15 December 1988). But soon after the award was announced, Sumitomo filed a claim for interference with a similar patent application it had filed within a day of the MIT researchers, according to Vear. Such claims of interference can be expected to multiply dramatically in the near future.

Patent interference cases in the United States can take years to resolve. One example is the battle over patent applications for the gas discharge laser, which took 28 years. The dispute over patent rights on high-temperature superconductor materials may become such a case because of the large number of parties involved. But a long legal battle could work to the eventual winner's advantage. If, as many experts predict, commercial applications of the new materials will not be feasible for at least a decade and because, in the United States, the life of a patent begins only after it is issued, the group to which the patent is eventually granted may enjoy licence fees for longer than if its claim had never been challenged.

David Swinbanks

■ On page 49 of this issue, the recently developed 'electron' superconductors show behaviour similar to the more established 'hole' superconductors. □

Personal communication

London

THE Union of Scientists and Inventors' Societies of the USSR has established a special "international commission for relations with former compatriots", aimed at establishing links with Soviet scientists and engineers now working in the West. Such people have hitherto been regarded as ungrateful renegades, who have accepted the benefits of a Soviet education and then decamped to the capitalists. The new commission, however, considers them a potential channel for establishing closer cooperation between Soviet science and the West. Interviewed by *Sotsialisticheskaya Industriya*, commission member Academician Yuri Gulyaev suggested a whole range of ways in which the *émigrés* could help Soviet science, from taking Soviet graduate students into their laboratories for a year or two to organizing joint research groups. Many countries, Gulyaev said, gained considerably by working with their *émigré* scientists in this way — Hungary especially. "We have been accustomed to working only through official channels" . . . but personal contacts and informal links are sometimes more effective." V.R.

Three per cent is six

London

BRITISH universities will receive a 3 per cent increase in government grants for recurrent expenditure in 1989–90. The minister for higher education, Robert Jackson, announced that £1,680 million will be distributed to the universities in the next academic year, an increase of 6 per cent from 1988–89, but much of the increase included in the government figures is accounted for by costs for restructuring and agreed pay awards for previous years. The universities' vice-chancellors say the 3 per cent increase is too small, and have asked for an increase of 5 per cent. Jackson also said the universities should plan for an increase in 1990–91 of 4 per cent in cash terms and increases of 3 per cent in the following two years. C. McG.

Chip complex destroyed

New Delhi

INDIA'S microelectronics industry has been shaken by a fire that almost totally destroyed the \$100-million semiconductor complex in Chandigarh, the country's only facility for producing silicon chips for civilian needs. Sabotage has not been ruled out as the cause of the fire, which burnt down the entire complex, leaving only some design documents and the administration block. Mr Virendra Mohan, chief of the facility, has resigned after acknowledging moral responsibility for the accident. There are now fears that many of the nearly 200 engineers working at the Chandigarh complex and who are now out of work may migrate to the West.

India's semiconductor needs will now have to be met by imports or by setting up a new facility that will not be ready before 1992.

K. S. J.

Oceanographers' lobby

Washington

To help convince Congress that oceanographers should play an important role in plans to address the issue of global change (see page 3), representatives from some four dozen oceanographic institutions met last week in Washington to form the Council on Ocean Affairs.

Jim Baker, president of the Joint Oceanographic Institutions, which has sponsored the new council, says that while federal agencies such as the NOAA and NASA are knowledgeable about the needs of oceanographers, Congress needs independently to be educated about the issues important to them.

While the council is intended to represent institutions, the new Oceanography Society, also promoted by Baker, is a professional society for oceanographers. The society will hold its first annual meeting this summer in Monterey, California. J.P.

Staab steps down

Munich

HEINZ Staab, president of West Germany's Max Planck Society (MPS), has announced that he will resign when his six-year term ends in the summer of 1990. A newly formed selection committee will present a list of potential replacements to the MPS senate in November.

Staab, 62, who is also a director of the Max Planck Institute for Medical Research at Heidelberg, would have been eligible for one more term. He is stepping down to devote himself to research. S.D.

Melbourne rivalled

Sydney

SYDNEY may be about to take a step to right the balance between itself and Melbourne, traditionally the centre of Australian medical research.

Professor Tony Basten has been appointed director of the new Centenary Institute of Cancer Studies and Cell Biology, a joint venture between the Royal Prince Alfred Hospital and the University of Sydney. The new institute will incorporate the Clinical Immunology Research Centre at the University of Sydney and is due to move a staff of 150 to two floors of a new six-storey building by October 1990.

The institute will be supported by state and federal grants, a block grant from the National Health and Medical Research Council and money from the private sector. T.E.

Genentech's heart drug

Washington

THE US Food and Drug Administration ruled on Monday of this week that Genentech can claim in its advertising that TPA (tissue plasminogen activator) reduces the death rate among patients experiencing a heart attack, based on new data from two ongoing clinical trials comparing several heart drugs. Previously, the agency would only allow that TPA treatment dissolves the blood clots causing the heart attack. C.E.

Freeze hits BMFT research

Munich

BATTLE lines have been drawn for the fight for next year's budget for the West German Research and Technology Ministry (BMFT), with European space projects, microelectronics and other politically sensitive items hanging in the balance. The federal research minister, Heinz Riesenhuber (Christian Democrat or CDU), has even come under attack from his own party as well as the opposition for not fighting harder for an increase in his stagnating budget.

BMFT is responsible for the general promotion, planning and coordination of research in West Germany. It supports technological research and development, data processing, nuclear research and technology as well as space and aeronautical research. As well as providing the lion's share of support for the Large Research Establishments (*Grossforschungseinrichtungen*) and the Max Planck Society, BMFT supports basic research at universities in selected areas.

The BMFT budget will decrease in real terms in 1989 for the first time in recent memory, if a budget freeze imposed last November is not lifted. The ministry lost DM190 million of its DM7,600 million budget to a "global reduction" imposed on all ministries. A further DM166 million has been frozen by the finance minister. Coupled with the effects of inflation, these cuts result in a reduction of the overall amount spent by BMFT in 1989, instead of the 2.8 per cent increase announced last year.

Ministry officials are not optimistic that the freeze will be lifted, but are trying to estimate where the axe will fall. A likely candidate for a cut would be the West German contribution to the large projects of the European Space Agency (ESA). But ESA would have to approve any such reductions, the effect of which might be to delay projects such as the space shuttle Hermes or the booster rocket Ariane 5.

The Max Planck Society, 60 per cent of whose funds derive from BMFT, will be largely protected from the cuts, thanks to the efforts of the West German finance minister Gerhard Stoltenberg and the *Länder* (states).

The shape of the 1990 budget will emerge in the next few weeks from discussions between BMFT and the Finance Ministry. The government will present a budget to parliament only in July, after its broad scope has been decided behind closed doors.

Meanwhile, there are signs that the ministry will shake off its lethargy and, even at this stage, ask for a big increase in 1990, partly to make up for the cuts in

1989. A ministry spokesman would not say how big an increase Riesenhuber will seek, except to say that it will be more than the growth of the overall federal budget. Some predict that the minister will ask for an extra 6 per cent.

West German participation in Hermes and the space station Columbus has begun to look even less certain than before in the light of the struggle over the BMFT budget. The consensus in Bonn is that it will probably not matter if a few tens of millions of deutschmarks are delayed until next year. But a much bigger commitment will be necessary in 1991 if Hermes and Columbus are to get off the ground. So far, the money required, eventually amounting to thousands of millions of deutschmarks, has not shown up in the medium-term West German budget.

In a related development, the Bundestag budget committee voted last week to release a previously frozen DM27 million for supersonic transport technology. The money will be invested in developing the proposed West German space plane Sänger, which represents the next generation of space shuttle technology.

Yet West German research and development continues to grow. But figures made public by BMFT show that the rising investment in West German research and development derives primarily from industry. No less than 74 per cent of the

Heinz Riesenhuber (right) is under pressure from his own party and the opposition. Social Democrat member Josef Vosen complains "Riesenhuber has promised a lot, but the money is not [yet] there".



DM61,400 million spent in 1988 was industry money. But a BMFT official says it is "questionable" whether overall research spending will continue to grow as fast as the gross domestic product, given the slow growth in the BMFT budget and the loss in 1990 of tax deductions for some company research.

BMFT notes that the trend elsewhere, in Japan and the United States for example, is just the reverse, with governments spending "disproportionately much" on research. But a BMFT official adds that there has recently been a "storm" of grant applications from university researchers, suggesting that the stiff competition for grants experienced by the Deutsche Forschungsgemeinschaft (DFG) has now hit BMFT as well (see *Nature* 337, 590; 16 February 1989).

Steven Dickman

SWISS RESEARCH

OECD urges changes for 1992

Munich

SWEEPING changes in science and technology policy will be required if Switzerland is to meet the challenge of the internal European market planned for 1992. That is one conclusion of a report* released by the Paris-based Organisation for Economic Co-operation and Development (OECD) in Berne this week. The review was carried out at the request of the Swiss government.

The survey team of three international examiners based their report on a series of site visits to Swiss laboratories, universities and public and private institutions.

The recommendations urge changes in policy at federal and cantonal levels as well as in universities and in industry. The reviewers found a "distressing" lack of cohesion in national science policy, with a resulting parochialism and slowness to change. Such traits may leave Switzerland "almost isolated in Western Europe" given the country's decision to remain outside the European Community (EC).

The most striking recommendation is the creation of a "secretary of state" for science and technology in the Federal Department for Home Affairs. This official would coordinate federal and cantonal policies as well as Swiss participation in international programmes.

The OECD report encourages Switzerland to invest more in research and to do so more effectively. In spite of being the richest of all OECD member countries, the report says that "public resources for research do not match those of most other OECD countries". The high percentage of the gross domestic product invested in research comes about because of research-intensive pharmaceutical and electro-mechanical industries. OECD strongly recommends that the Swiss government streamline the diffuse structure for research policy-making and establish a "national research budget" — at present decisions are made at a local level in the departments.

The lack of major public research institutions (with the single exception of the Paul Scherrer Institute for nuclear research) comparable to those elsewhere in Europe gives the team the impression of "weakness, dispersal and fragmentation of public R&D funding," said OECD.

Among the new programmes recommended by OECD are "substantial" national research programmes in information technology, new materials and biotechnology. The lesson learned when Switzerland "missed the turning" on the electronics road — and failed to develop an indigenous industry in digital watches or other electronic products — must not

be forgotten, warned OECD.

OECD concurred with some university researchers and industrialists (see *Nature* 336, 331; 24 November 1988) that Switzerland needs a publicity campaign explaining the benefits of science to the public and to school leavers. Distressingly few Swiss students are studying science and engineering, raising doubts about the future supply of skilled personnel.

On the Swiss National Science Foundation, which supports much of the basic research at Swiss universities, the report called "unhealthy" the practice of reducing individual project allocations by up to 30 per cent to allow more proposals to succeed. This encourages groups to submit inflated grant proposals and can prevent some projects from achieving "critical mass."

RESEARCH COUNCILS

MRC offers more

London

THE British Medical Research Council (MRC) wants to increase the pay of graduate students in the hope of stemming the growing shortage of young researchers. Last week, the council asked the Department of Education and Science (DES) to allow an increase from just under £3,000 to £3,500 in the stipends paid to the holders of its basic research studentships from next September. In doing so, the council is planning to break with the traditional method of calculating graduate students' pay using a DES formula linking it with the maintenance grants for which undergraduates are eligible.

The council also wishes to offer an extra £500 a year to graduate students working on its directed AIDS programme of research, which would be the first time that pay differentials not based on location had been offered. The stipends paid to holders of research council studentships are usually tax-free.

The problem of attracting talented young researchers is also worrying the Advisory Board for the Research Councils. At a meeting organized by the Royal Society last week, ABRC chairman Sir David Phillips said that the shortage of research manpower would be the most serious problems the research councils would have to face in the 1990s.

ABRC's Biotechnology Advisory Group is now urging the board to set up a formal review of the manpower shortage and to suggest ways in which it might be remedied. One member of the group, Professor Peter Dunhill of University College, London, says that if nothing is done about the shortage, all the councils' initiatives in biotechnology could fall apart.

Christine McGourty

The report repeats the commonly voiced fear that, without subsidies, the small and medium-sized companies forming a large part of Swiss industry will not be able to compete with their rivals in EC countries.

Rounding out the OECD recommendations is a list of conditions to be met if Switzerland's "production apparatus" is to face the needs of modernization. The list was first prepared by the Battelle Institute in 1983 in a study of what had gone wrong with Swiss industry in the 1970s. The Battelle Institute said Swiss industry should take more risks in introducing new technology, that large and small companies should cooperate more closely, and that the banks should take a greater interest in the promotion of new technology. OECD said that the recommendations are "still valid today."

Steven Dickman

NAGYMAROS DAM

Objectors find a voice

London

THREE members of the Hungarian National Assembly are in danger of losing their jobs because of their electors' objections to the controversial Gaboikov-Nagymaros hydroelectric scheme. After a majority of the assembly voted for completion of the Hungarian part of the scheme, three constituencies (two in Budapest and one in Szeged) have voted to recall the mandates of their representatives under a previously unused provision of the Hungarian constitution.

The decisions reflect growing opposition in Hungary to the Nagymaros dam, part of a joint project with Czechoslovakia. There are demands that the government should annul its treaty obligations to Czechoslovakia (as well as the construction debt to Austria), and that Hungary should pull out of the scheme unilaterally.

The decisions also reflect a new political awareness in Hungary. In a one-party system (as at present), the right of recall is a guarantee that constituents' views will be accurately represented in the assembly. (A fourth deputy is threatened with recall for reasons unconnected with the dam.) But there are signs that the ruling Socialist Workers' Party may abdicate its monopoly; the Central Committee has now accepted the principle of a multi-party system with appropriate "checks and balances".

But whether this will allow the emergence of a 'green' party in Hungary is uncertain. Janos Vargha, one of the leaders of the 'Danube circle' which has campaigned against the dam told a recent conference of environmentalists that a green party could compromise the independence of existing groups.

Vera Rich

*Reviews of National Science and Technology Policies: Switzerland. The report, including an account of the review meeting held in Berne on 3 March, will appear in book form later this year.

European satellite reception

Paris

THERE is now hope that the European Remote Sensing Satellite, ERS-1, to be launched next year, will have an enthusiastic band of users. That seems to be the outcome of last month's meeting between the director-general of the European Space Agency (ESA), Reimar Lüst, and the President of the Commission of the European Communities (EC), Jacques Delors.

The satellite will allow Europe to have a major role in studying the world climate and the effects of pollution. But although the infrastructure of ground stations to receive data from ERS-1 is already in place, the network of potential users is still relatively underdeveloped but last month's meeting seems to have helped.

ERS-1 will also be at the heart of the World Ocean Circulation Experiment (WOCE), which will come alive, after 10 years of preparation, with the satellite's launch. Its advanced radar instrumentation will allow accurate data to be collected on sea surface temperatures, ice flows and ocean circulation, as part of an international effort to understand global

climate and natural disasters. The satellite could also be used to monitor the rate of destruction of tropical rain forests, often invisible to other remote-sensing satellites because of cloud cover.

The meeting between Delors and Lüst was aimed at soldering links between ESA and the European Community. "The European Commission is like a dream client", said Jean Arets, head of international relations at ESA. One possibility seems to be a high-level working group to encourage environmental monitoring by satellite. Arets also says that the European Community could help to finance the creation of a community of users.

The goals of ESA and the EC are reciprocal. ESA is keen to improve its coordination with the EC's new technology programmes, while the EC wishes to strengthen pan-European involvement in ESA programmes for Earth observation. It also seeks a long-term strategy to exploit microgravity techniques, and has an even more immediate need of common norms for telecommunications.

Peter Coles

AURORAL SATELLITE

Japan continues in an old furrow

Tokyo

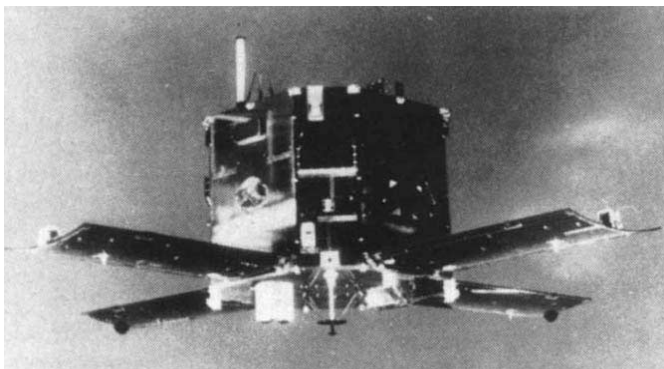
THE first satellite designed specifically for study of the particle acceleration mechanisms producing the aurora borealis (and australis) was launched successfully last week (22 February) by Japan's Institute of Space and Astronautical Science (ISAS) from the Kagoshima Space Center.

The EXOS-D satellite was carried into a polar orbit by a Mu-3S II three-stage solid-fuel rocket, independently developed under ISAS's scientific space research programme. The 300-kg satellite, given the post-launch name of 'Akebono' (Dawn), will fly into the region where charged particles from the solar wind are accelerated by the Earth's magnetic field and penetrate to the upper atmosphere, causing auroral displays. The orbit will take the satellite from a near point of 300 km from Earth out to 10,000 km.

On board Akebono are instruments to monitor magnetic and electric fields, plasma waves, low-energy particles, supra-thermal ions and thermal electrons as well as a special auroral camera

designed to take photographs in the visual and ultraviolet spectrum. Early next year, ISAS hopes to add to Akebono's view by launching sounding rockets capable of reaching a height of 350 kilometres from a base in Norway, making it possible to obtain the first simultaneous observations of the aurora from above and below.

Akebono is ISAS's 12th scientific observation satellite and the third in a series designed to look at the aurora and



Akebono: a first in aurora research.

magnetosphere, following on from Kyokko ('Northern lights') and Jikiken ('Magnetosphere'). ISAS plans to continue providing a key role for Japanese researchers in solar wind and magnetospheric research with the launch of a satellite into the Earth's magnetotail in 1992 as a part of the International Solar Terrestrial Physics programme. Alun Anderson

Close call in election fight

Tokyo

A TOUGH election battle for the presidency of the University of Tokyo resulted in victory last week for Professor Akito Arima, one of the chief proponents of major change for Japan's leading university.

The election was so close that the university was eventually compelled to choose between two candidates by asking them to draw straws (actually a pair of bamboo sticks with concealed, differently coloured tips). It is the first time that this chance procedure has had to be used in the 112-year history of the university.

Arima (58), a theoretical nuclear physicist, will become president of the university for a four-year term beginning on 1 April. He is the first president to come from the science faculty for more than thirty years and faces a university which is divided over the plans for radical educational reform led by the science faculty.

The science faculty's plans involve fusing undergraduate and graduate schools into a new six-year school (see *Nature* 330, 597; 17 December 1987). But opponents fear it may lead to a drop in prestige for the College of General Education, whose faculty provides the first two years of education to all students and less-specialized further education in departments of Liberal Arts and Pure and Applied Sciences.

Department of Liberal Arts' Professor Nagoya Homma, who has opposed the reform plan, proved to be Arima's chief opponent.

To begin with, there were five candidates and after three rounds of voting the two top candidates both drew exactly 586 votes from the assembled professors, associate professors and lecturers of the university.

At that point, previously unused university rules specified a decision by lottery and Arima drew the winning red-tipped stick.

"Exhausting" was how one Senate member described the day-long election. But a much longer campaign lies ahead for those favouring reform. The members of the science faculty have seen top university research institutes turn into independent national facilities with excellent equipment and have seen private companies set up basic research institutes with large financial resources. Without some strengthening of the university research system, the science faculty are apprehensive that they will find themselves seen as a source of trained students rather than as leading basic researchers.

Alun Anderson

COMPUTER VIRUSES

No immunity for Moscow

Moscow

COMPUTERS in the Soviet Union have been suffering from virus infections just as in the United States and Western Europe. In one of the most recent cases, at the Institute of Applied Mathematics of the Soviet Academy of Sciences, researchers found that some of their programs were disappearing for no apparent reason shortly after retrieval. Their computers had been attacked by a virus. The rogue program was quickly found and killed.

Virus infection is not new. About 20 years ago, programmers at a research institute were annoyed by the erratic behaviour of an ICL computer. The institute was on the point of sending for help from the British supplier when a young programmer admitted having written a program that fooled the operating system while remaining invisible.

With the advent of personal computers, a little too late in the Soviet Union, programmers have been confronted with a formidable job of developing anti-viral codes. The first Soviet machines to suffer from viral infections, ostensibly imported from the West, were Atari computers, followed by IBM-compatible machines.

By the time the International Computer Club was formed in Moscow last December, Soviet programmers had compiled a list of lethal files, including viruses and so-called "Trojan" programs. A user with such a list can identify the infection but cannot cure it. It takes a special anti-viral program to purge the virus.

There are now some 40 virus programs on the Soviet danger list and the number continues to grow. The computer epidemic is spreading because of the woefully inadequate software distribution system in the Soviet Union — or rather because of the lack of any such system. The result is that most Soviet programmers are engaged in entirely disorganized computer activities, seeking casual contacts with colleagues.

With no official software market, such contacts are the only way to obtain software in high demand. Among several hundred of my acquaintances who work with software best-sellers, only a handful have bought them legally.

There is no evidence that specific Soviet computer viruses have been, or are being, developed. Indeed, the best Soviet programmers are said to be working on a computer 'vaccine'. Such a program was described by the central Soviet newspapers last autumn. The idea is that the program identifies a constant part of a known virus and removes it from the host program. But such programs are not universal, as they purge only well-known viruses with 'recognition sites'.

Apart from rudimentary restriction programs, there are two anti-viral 'vaccines' in the Soviet Union. The first is a resident program monitoring a computer's 'abnormal' behaviour, which may suggest that a virus attack is under way. The other may seem more attractive to users in that it is not memory-resident and may be run after a series of operations to give a programmer a list of files which may have been modified by a virus.

Having analysed the suspicious programs the programmers can tell whether the computer system has been infected. One difficulty is that the incubation periods may be very long so that the owner of the damaged machine may discover a new infection.

Worse still, until recently there was no popular magazine for programmers, so that news about viral attacks was spread by word of mouth. The most lethal viruses, which are spread by networks, pose no danger to Soviet computers; there are almost no computer networks in the Soviet Union. Meanwhile, the need for computer virologists and for joint efforts with Western colleagues is increasingly evident.

Pavel Pevzner
Novosti

ANTARCTICA

Greenpeace's claims refuted

Paris

THE director of the research mission of the French department for Austral and Antarctic Territories (TAAF) has rejected claims by the environmental group Greenpeace that an airstrip being constructed in the French Antarctic Territory contravenes the Antarctic Treaty (see *Nature* 337, 106; 1988). Bernard Morlet says the airstrip is being built to support scientific research, and that steps are being taken to ensure that colonies of Emperor and Adélie penguins are not endangered by construction work.

Plans to build an airstrip at Dumont d'Urville date back to 1983, and are prompted by seasonal restrictions on sea access to the French base, which is surrounded by several hundred square miles of pack-ice in winter. Ships can first reach the base at mid-summer in December, and researchers arriving in January, when the base has been made habitable, have only a month to collect data if they are to avoid being icebound. An airstrip, says Morlet, will allow researchers to winter at the base and to begin work — in human and animal physiology, glaciology and climatology — once summer begins.

According to Greenpeace, the airstrip has destroyed breeding grounds for a colony of Adélie penguins and blocked the migratory path of Emperor penguins. But Morlet says that the migratory path of the Emperor penguins is variable, so that the construction may pose no problem at all, and that, in any case, they are used to climbing onto sea islands to avoid winter storms and so should be able to surmount the airstrip, which is only 5 metres above sea level.

Morlet does acknowledge that the breeding ground of the Adélie penguin, jeopardized by the landing strip, is a problem, but says that the effect of the construction project can be minimized by protecting the adult population and encouraging a change in breeding. He says that the material used for the airstrip may help to create new breeding sites. He also accuses Greenpeace of misreading the Antarctic Treaty. While its third (1983) revision forbids harmful interference with bird and seal colonies — by the use of explosives or by disturbance during their breeding periods — it also allows these activities "...to the minimum extent necessary for the establishment, supply and operation of stations".

Morlet also rejects Greenpeace's allegation that the airstrip is a foot in the door for mineral exploitation. He says that the 2,000-metre icepack beneath the French base makes it unsuitable for this purpose.

Peter Coles

Gorbachev's first Chernobyl visit since disaster

Associated Press



Mr Gorbachev and his wife Raisa pictured talking to workers during a visit to the Chernobyl nuclear plant last week to inspect the site of the world's worst nuclear disaster three years ago. Mr Gorbachev promised that nuclear power, while essential, was unthinkable without guarantees of complete safety — a sensitive issue in the Ukraine, which has the greatest concentration of nuclear power stations in the Soviet Union.

Human space flight

SIR—The maxim quoted again by G. B. Field, M. J. Rees and D. N. Spergel in their otherwise excellent article (*Nature* 336, 725–726; 1988) that human space-flights “cannot be expected to yield a return commensurate with their cost if the judgement is made in purely scientific terms” is objectionable.

We need look no further than the Viking missions to Mars to see the problems of trying to automate science. Even after the expenditure of \$1,000 million (more than \$3,000 million in current dollars), nobody knows today whether or not there is life on Mars. The Viking landers failed in the single purpose for which they were built. This is not to say that they did not serve other ancillary functions, but had the landers not been constructed around elaborate biology packages, the initial reconnaissance of Mars and its system could have been accomplished for a small fraction of what was actually spent.

We forget that the Viking missions were not inexpensive relative to manned flights. Typical modern estimates for the cost of sending people to Mars (which are little changed from those during the mid-1970s) range from an unrealistic low of \$10,000 million to a more likely \$40,000 million — only 3.3 to 13.3 times the \$3,000 million cost of Viking in current dollars. Can anyone seriously argue that putting a few intelligent and flexible human beings and their equipment on the martian surface for weeks (or for months or years) will result in only thirteen times Viking’s scientific return?

Given the decisions that face the US space programme, a more relevant way to look at this issue is to consider that, even given the highest estimates for the true cost of launching a Spacelab shuttle mission (about \$500 million), the Viking expenditure would pay for no less than six Spacelab flights. Whichever way we measure it, even one Spacelab flight achieves far more science than did Viking. Does it not make sense to consolidate our position in the Earth–Moon system before we go gallivanting across the Solar System? Perhaps this could be done by building the much-maligned space station, which, in addition to laying the technological foundations for human planetary exploration, would admirably serve almost every field of space science but planetary exploration. The space station would extend all the advantages of Spacelab at a cost of only some eight times Viking.

The distinction between reconnaissance and science may seem fine, but I believe it is the natural division between what should be expected of automated spacecraft and what must wait for human

exploration. The lesson of Viking is that, for the foreseeable future, automating creative science will at best be extremely expensive and difficult, and probably impossible — whether on Earth, in orbit or far across the Solar System. This is a lesson we should consider very carefully before spending vast sums of money (\$16,000–10,000 million is the current guess) on a Mars Rover which roves less far with every study, trying to automate what a human driver could do far better for only a few times the cost.

Automated missions are tremendously valuable for some limited applications. But human missions are necessary for most sciences beyond initial reconnaissance, and planetary scientists do themselves no favours by denying that.

DONALD F. ROBERTSON

940 Capp Street,
San Francisco, California 94110, USA

The test of time

SIR—I have wasted a lot of time in my life trying to reproduce astonishing results. I know of scientists who, on any subject, publish in parallel two or three competing interpretations or theories. And a few years ago I was amazed to hear a scientist prominent in my field disclaim five articles he had recently published. It is difficult to convey the resulting scepticism about the literature to young scientists without discouraging them.

I therefore propose that when a scientist is under consideration for a new post or for promotion, his or her publications list should be checked to see whether the findings and interpretations have stood the test of time and, if not, whether they have been retracted. This complements the Harvard rule whereby only positive evidence (the four best papers) is considered. This would uncover quite a few prominent researchers whose latest sensational work always seems to obscure past errors.

JACQUES E. DUMONT

Université Libre de Bruxelles,
Faculté de Médecine,
Campus Hôpital Erasme,
Route de Lennik, 808,
B-1070 Brussels, Belgium

Rough justice

SIR—William McBride has some valid comments in his letter on the inquiry into his alleged fraudulent conduct (*Nature* 336, 614; 1988). On the basis of his statement that he was not allowed legal representation, nor to attend all the hearings, nor to cross-examine witnesses, he did not receive natural justice.

The transcript of the inquiry is not publicly available and analysts have only

the report to guide them. It seems that McBride should perhaps have been given the benefit of the doubt on having altered the oral dose rates because he believed he was correcting them in answer to a referee’s comment. The report makes a convincing case for finding that McBride did not use “proper scientific method” (*Report* 25; 1988) in his work. This might explain his dubious method of corrections and his incorrect statements about controls. It seems, however, that there was “deliberate falsification” (*Report* 24; 1988) about the number of experimental rabbits.

The final version of McBride’s paper, published by an Australian journal after its rejection by an overseas journal, should have been scrutinized more carefully by the referee in order to remove unsupported conclusions. Foundation 41 and its research advisory committee also performed badly, especially in their failure to investigate fully the staff allegations and to oversee remedial action.

G. F. HUMPHREY

School of Biological Sciences,
University of Sydney,
Australia 2006

Apoptosis

SIR—In reply to B. Kleine (*Nature* 377, 402; 1989), J.F.R. Kerr *et al.* (*Br. J. Cancer* 26, 239–257; 1972) proposed the term “apoptosis” (suggested by Professor J. Cormack, Department of Greek, University of Aberdeen) for a process of active, programmed cell death with a number of characteristic features which clearly distinguish it from “necrosis”, passive cell death. The standard Classical Greek–English dictionary (H. G. Liddell & R. Scott, *A Greek–English Lexicon*, 6th Edn, Oxford University Press; 1869) defines ἀποπτῶσις as “a falling off or away”; it is used of the falling of leaves, a clear example of programmed death. The cognate verb is, of course, the reduplicated form ἀποπίπτω; ἀποπτῶσω, quoted by B. Kleine, is not known to this dictionary. A computer search of papers published since 1983 revealed fifty with “apoptosis” or “apoptotic” in the title. Had B. Kleine consulted A. H. Wylie *et al.* (*J. Pathol.* 142, 67–77; 1984), cited in our letter (*Nature* 337, 181–184; 1989), rather than his Greek–German dictionary, he would have found a clear exposition of the meaning of apoptosis.

CHRISTOPHER A. SMITH

GWYN T. WILLIAMS

ROSETTA KINGSTON

ERIC J. JENKINSON

JOHN J. T. OWEN

Department of Anatomy,
The Medical School,
University of Birmingham,
Birmingham, B15 2TJ, UK

Pan-African ornithology divided

T. M. Crowe

A congress that once offered a forum for all African ornithologists has been dealt a crippling blow by politics. African birds and those who study them are the losers.

THE first Pan-African Ornithological Congress (PAOC) was held in Livingstone, Northern Rhodesia (now Zambia) in July 1957. The concept of the congress was the brainchild of a South African, Dr Cecily K. Mackie Niven, who was stimulated to this end chiefly by politically motivated rejections of her invitations, on behalf of the South African Ornithological Society (SAOS), to host the International Ornithological Congress in southern Africa. The first PAOC, with Niven as president, and the next two, were held under the auspices of the SAOS; the editors and many of the officials of those and its two succeeding congresses were South Africans. Dr Gerard J. Morel (chairman of the sixth congress) acknowledged that PAOC is very much a South African invention: "The credit for this unique and enviable situation [the PAOC] is largely to be attributed to our colleagues of the Southern African Ornithological Society".

In 1985, delegates at the sixth congress, held in Francistown, Botswana, voted overwhelmingly for Nairobi, Kenya, as the venue for the seventh congress in August 1988, on the assurance given by Kenyan delegates that a significant number of South Africans would be granted visas to attend. But by 1987 it had become apparent that, for political reasons, South African participation would be at best token. Indeed, for a time there was some doubt as to whether the meeting would take place at all; the Kenyan government officially approved of the congress only a few months before it was due to be held.

Fearing the consequences of a PAOC without contributions from South African ornithologists, many of the members of the PAOC *ad hoc* committee asked the local organizing committee in Nairobi to obtain an unambiguous statement of the Kenyan government's policy on the admission of South African delegates. Some also asked fellow committee members to consider an alternative venue if delegates should be denied visas. Indeed, Nairobi had been abandoned as a venue for two previous congresses on just these grounds.

The response to the first request was vague and non-committal, but it seemed that the Kenyan government would almost certainly not issue visas to South

African passport holders. The response to the second request was that the congress would be held in Nairobi "with or without South Africans". Thus, neither the *ad hoc* nor the local organizing committees considered the presence of South African ornithologists to be essential; prospective participants would have to apply individually for visas and would simply have to "take their chances".

"One can only speculate as to how much the scientific quality of the congress could have been improved by the contributions of South African delegates, not to mention the thousands of dollars in registration fees that would have been paid by dedicated amateur South African ornithologists."

In January 1988, those few South African ornithologists who still hoped to attend in August were instructed to provide passport details to the local organizing committee, which would then submit a block application to the Kenyan immigration department. About 15 of us provided these details. But, about three weeks before the start of the congress, I received by courier a batch of visa application forms which had to be filled out individually and returned within that week. On the same day, I received a cable notifying us that our chances of receiving visas had been "weakened by shortage of info from SA". By this time, only six of us still persisted in our determination to attend the meeting.

On the afternoon of 23 August 1988 (five days before the start of the meeting) I was told by telephone by Professor G. M. O. Maloiy, the conference chairman, that my application for a temporary visa had been approved, and that my visa would be in the hands of immigration authorities at Nairobi airport. I was also notified that a similar application for Richard K. Brooke, my colleague at the FitzPatrick Institute, had been approved, but that those of Ian A. W. Macdonald, another institute colleague, Professor Eric H. Harley (of the Department of Chemical Pathology, University of Cape Town) and of Professor Leslie G. Underhill (of the university's Mathematical Statistics Department) and his wife "... were still under consideration". Unfortunately, Brooke and Macdonald had by then cancelled their plans to attend, and Underhill

and his wife were notified by telephone on 26 August (and Macdonald subsequently by post) that their visa applications had been refused. Harley was informed of the success of his visa application on 30 August (two days after the start of the Congress).

Thus, only Harley (holder of a British passport) and myself (a US passport holder) actually attended the seventh congress, as 'official' South African delegates. Dr Richard Liversidge, also a South African resident, had travelled via London on a 'clean' British passport (and thus did not require a visa).

I arrived at Nairobi airport on the evening of 27 August to find that the Kenyan immigration authorities had no record of my visa. Fortunately Dr Hussein Isack, chairman of the local organizing committee, arrived in the nick of time with a photocopy of the visa, and kindly drove me to the hotel into which I was booked. After one evening's stay, I concluded that the hotel was not suitable: two delegates had had their rooms rifled and money and passports stolen; another delegate was brutally mugged nearby. Moreover, the hotel was frequented by aggressive prostitutes and was a costly cab ride from the congress venue, the National Museum of Kenya. Therefore on 28 August I shifted to another hotel which, although it was a mere five-minute walk from the museum, had not been offered as possible alternative accommodation by the local organizing committee.

When I registered that same afternoon I found that my delegate's badge listed the United States as my country of representation (Liversidge and Harley were listed as from the United Kingdom); the country name had been deleted from the abstracts of the two papers I was to present. My packet of congress materials also included a letter from the chairman of the *ad hoc* committee threatening to disrupt the presentation of my papers and to prevent me from attending future congresses if I did not submit manuscripts for the congress proceedings during, or shortly after, the meeting. (I had questioned the hard-and-fast rule in the proposed PAOC constitution that no oral paper could be presented if a complete manuscript was not also delivered.)

Although the seventh congress was

well-attended (more than 100 papers and more than 200 delegates from 19 African and 14 non-African countries), the organization of the programme, the quality of the papers and their relevance to African conservation biology left much to be desired. The original intention had been to have sessions or mini-symposia with themes (seasonality, birds as biological indicators, nest parasitism, pest birds, biogeography and systematics, and so on), many of these were decimated by the enforced withdrawal or exclusion of South African ornithologists. For example, my session on systematics and biogeography had the heart ripped out of it with the absence of Dr Alan C. Kemp (Transvaal Museum), Tony Harris (Transvaal Museum) and Dr W. Stewart Grant (University of the Witwatersrand). I read brief summaries of Kemp's and Grant's papers, but papers delivered *in absentia* have little impact and do not appear in the congress proceedings.

Similarly, the session on birds as indicators of environmental change, which had been organized by Ian Macdonald, was, without him, a ghost of what it could have been. One can only speculate as to how much the scientific quality of the congress could have been improved by the contributions of South African delegates, not to mention the thousands of dollars in registration fees that would have been paid by dedicated amateur South African ornithologists.

On a more positive note, about a quarter of the oral papers were presented by indigenous Africans. But with certain noteworthy exceptions, their presentations were either dry or emotive descriptive accounts or 'how-to-kill-quelea'-type recipes. Few of these papers were based on statistically sound and analysed quantitative information. These deficiencies were certainly not due to the presenters' difficulty with the English language or a lack of dedication, enthusiasm or intellectual ability, but rather can be attributed to poor training in fundamental biological concepts and analytical skills.

Poor science

Perhaps the most disturbing aspect of the meeting was the poor scientific quality of the conservation papers. Clearly, the main threats to African birds relate to the fragmentation of their habitats (because of man's land-use practices) and concomitant reduction in species' populations. It is simply not good enough to preserve a patch of the preferred habitat of a species if that is too small and isolated to maintain a minimum viable population. Only one paper, by W.D. Newmark, and delegates at a workshop on forest ecosystems convened by Dr Simon Stuart, specifically addressed these issues. At a similar workshop on savannah ecosystems, the convenor was apparently unaware of

the importance of studies of minimum viable population size.

Papers presented during the 'conservation' day devoted to the International Council for Bird Preservation were mainly conservation status-type papers and accounts of educational programmes. There was little or no quantitative biogeography, ecology, genetics or socio-economics. Our job as ornithologists is to provide conservationist organizations' negotiators with the necessary scientific ammunition to sway decision-makers in governments. It is not sufficient merely to make resolutions that this or that chunk of habitat needs to be preserved, but rather to make explicit, scientifically sound predictions about the consequences of its preservation or destruction.

The key clauses in the PAOC constitution that will effectively prevent South African participation in the two remaining congresses this century are the aim to rotate the venues among the four regions of Africa (east, west, central and southern) and the limitation of a single country's membership on the PAOC committee, the ruling body, to a maximum of 3 out of the 20 members. The selection of venues means that the next two congresses will almost certainly be held in central and west Africa. The apportionment of committee members on a strictly regional basis takes no account of geographical variation in the abundance or quality of African ornithologists. Thus, in theory, there could be no South African representation on the committee, but west Africa, which appears now to have no trained resident scientists at PhD level working full-time on birds, will have four members on the committee. When I brought up the issue of the importance of free access to all in choosing a PAOC venue, only one committee member felt that this should even be considered.

One positive development was nevertheless that the committee agreed to consider being indirectly affiliated with the International Council of Scientific Unions. If PAOC adopts the council's code, no venue could be chosen which would exclude delegates on political grounds. But this would not help amateur South African ornithologists, most of whom would still be unable to attend as they do not present papers, the rules applying only to delegates who do present them. This is a particularly tragic blow to African ornithology as nearly half of the delegates attending the first six PAOCs were South African, most of them being dedicated amateurs.

The primary criterion which influenced the choice of the venue for the next PAOC was that it must be in a French-speaking African country. Other factors, such as ease of access (financial and political), excellence of site and excursions or the relevance of birds to local conservation,

were of secondary importance. On this basis Kigali, Rwanda, was chosen as the venue of the eighth PAOC to be held in 1992, with Ghana as a prospective venue for 1996.

Bitter pill

The PAOC created by Niven is now extinct, and has been replaced by a very different sort of animal. I have reservations about the 'fitness' of this new creature, but will support it as long as there is a reasonable chance of it promoting African ornithology. My message to my South African colleagues is perhaps a bitter pill. As African ornithologists, however, we must endeavour to maintain our tenuous ties with our colleagues to the north. These people need and want our collaboration and we would be remiss if we allowed political differences between governments to sever our ties with them. The virulent anti-South African elements in the PAOC committee are but a comparatively small section of the tail of this animal, powerful enough at the moment to wag the dog, but we have every reason to hope that that power will wane.

At the same time, although the passing of the SAOS-style PAOC is to be lamented, we must continue to promote South African ornithology by sending our best ornithologists, especially young graduates, to discuss their research at international conferences. We must also promote ornithology locally by hosting international conferences in South Africa.

Finally, for too long ornithology in South Africa has been effectively a reserved white profession. This country and its wildlife belong to black as well as white South Africans. As a result of setting and maintaining high standards in our bird research, we have made it very difficult for educationally deprived black people to establish themselves as ornithologists. I believe that we (the SAOS and the FitzPatrick Institute) should do much more to publicize our activities and create incentives to promote the development of black ornithologists, and that we should take active steps to interest black South Africans in conservation biology and ornithology. The University of Cape Town (as well as other open universities) has active academic support programmes which are helping to redress the educational imbalance created by apartheid. □

T.M. Crowe is at the FitzPatrick Institute, University of Cape Town, Rondebosch 7700, South Africa.

ACKNOWLEDGEMENTS. I thank the Bremner Travel Grants Committee and FitzPatrick Institute (University of Cape Town) and the Foundation for Research Development of the Council for Scientific and Industrial Research for financial and other support that made my trip possible and successful. I thank Professor G.M.O. Maloiy and Dr Hussein Isack for their help with my obtaining a visitor's visa for Kenya and their friendship and hospitality during my stay. This article is a personal assessment of the seventh PAOC and does not necessarily reflect the views of my colleagues or the SAOS council.

How to say sorry graciously

Mistakes will happen, so that retractions will continue to be necessary. Two physicists have now produced a model of how the task should be tackled.

RETRACTION has become a loaded word, which journals and their contributors alike abhor. As things are now in the United States, there is even a danger that the mere appearance of a retraction will be regarded by a congressional committee or by a self-appointed watchdog as a signal for an inquiry into what has been going on. That is one reason why there should be some rejoicing that it is still possible to publish a full, even a fulsome retraction without bringing down the house on one's head. What has happened is that two physicists at the University of Maryland at College Park (almost in suburban Washington) have retracted the conclusions of an article that appeared last year in no less a journal than *Physical Review Letters* (59, 2507; 1988), but they have done so in such an open fashion that even their sternest critics will be disarmed.

The tenor of the retraction can best be judged from the sentence towards the end in which the authors, O.W. Greenberg and R.N. Mohapatra, say: "We are grateful to Rudolf Haag for warning us of the error of our ways". They then go on to thank another colleague for drawing their attention to an article in the Soviet literature which, if read in time, would also have warned them off. They acknowledge in their retraction (*Phys. Rev. Lett.* 62, 712; 1989) that many people have already embarked on experiments to test their earlier theory. But the best guess seems to be that those who have started down that road will continue, if only because the issue is in itself intrinsically interesting.

The question is whether there can be violations of Pauli's exclusion principle which, applied to electrons, forbids the simultaneous presence of two particles in the same state. An entirely equivalent way of putting this restriction is to say that the complete wavefunction of a system of several electrons (or other particles of the family called fermions) must be anti-symmetric with respect to the exchange of any two of them.

But that formulation has the advantage of suggesting that the wavefunction of a system of several bosons must be symmetric with respect to the exchange of random pairs of particles. The practical consequences are of course considerable. Fermions obey Fermi statistics and bosons (such as photons) obey Bose-Einstein statistics, which explains why the free electrons in a metallic conductor make

only a negligible contribution to its specific heat.

There is a third formulation of Pauli's principle which field theorists find more useful. A particle field, say some arbitrary distribution of radiation (photons), can be conveniently represented with reference to the dynamically possible states of a single photon in the available space. The simplest way of cataloguing the states is by means of the multiply infinite set of states defined by the standing waves which can exist within whatever space is accessible, each of which corresponds to a simple harmonic oscillator whose frequency corresponds to that of the radiation (when the bosons are photons).

Starting from that point and the principle that, in quantum mechanics, observable quantities are operators, the field theorists have fashioned a formalism in which the actual state of a system, which should be a specification of the numbers of particles in each of the possible oscillatory states, can be built up from the interaction of elementary operators, one set for each possible state, which have the effect either of creating or getting rid of a particle from that state. The creation and annihilation operators are the life-blood of field theory.

They are also the simplest means by which departures from Pauli's principle can be handled. For it is clear that the creation operators for fermion and boson fields must have very different properties. Two fermions cannot be in the same state, which means that although the effect of a creation operator on an empty fermion state is to fill it with a single particle, Pauli's principle implies that a second operation on the same state must be a nonsense. But with bosons, the creation operators must plainly behave differently, because there can be many particles in the same state. The algebra that arises by forming the products of several of these operators is simple but intriguing.

Greenberg and Mohapatra's original goal was to study small departures from Pauli's principle. Large departures would not have been surprising in the sense that there is a theory (going back to the 1950s) allowing for particles (which apparently do not exist) with properties intermediate between bosons and fermions, and called parafermions and parabosons. And, of course, they suggested experiments to tell how big the

departures are, or at least to define bounds for them.

Designing experiments is not as difficult as it may seem. If, for example, there is a chance that a single fermion state may occasionally hold two particles, it should be possible to discover that X-rays emerge from conductors carrying electric current, as electrons are occasionally captured into inner energy levels which temporarily violate Pauli's principle. Greenberg and Mohapatra report how some colleagues have completed such an experiment at Fermilab without finding anything untoward.

So where did Greenberg and Mohapatra go astray? Their starting point had been a modification of the algebra of the creation and annihilation operators, which they assumed to be feasible because it had been tried before. What they did not know was that a Soviet researcher, A.B. Gorkov from the international centre at Dubna, had shown as early as 1983 that the mathematical properties of the creation and annihilation operators forbid modifications other than those leading to the uninteresting states of parafermions and parabosons.

One virtue of this retraction is that the reasons why the argument went awry are fully explained, at least in language that those working in the field will readily understand. Another is the good humour of the piece, typified by the acknowledgment of Rudolf Haag "for having shown us the error of our ways". Yet another is the way that those who have helped to put the authors back on the straight and narrow path of rectitude are thanked for their help. And, finally, there are the grant-making agencies: support from the National Science Foundation is acknowledged in exactly the same terms in the two papers.

More than all this, the authors go on to give general reasons why their search for a basis on which Pauli's principle may be violated was a wild-goose chase from the start. The formalisms that give parabosons and parafermions correspond to the orthogonal and unitary symmetry groups of particular integral dimensions, but nobody has yet found a way in which these groups can be generalized into non-integral dimensionality. It seems rather hard on the authors to suggest that their retraction should be taken as a model in this rare genre, but that is what it is.

John Maddox

Initiation of recombination

J.R.S. Fincham and P. Oliver

RECOMBINATION within genes is generally by conversion — that is, by unilateral transfer of information from one allele to the other. It can be, and often is, accompanied by nearby crossing over, but it is most often the conversion rather than any correlated crossover that is responsible for the recombination within the gene. Much of what is known about this process comes from yeast and other fungi, where conversion shows up as non-mendelian ratios — 3:1 or 1:3 instead of 2:2 — in meiotic tetrads segregating with respect to single mutational differences. The complexity of the alternative models for the mechanism of recombination, although plausible, has hitherto detracted from their testability. Now that Szostak and colleagues have broken into meiotic recombination at the molecular level, as they report in two papers on pages 35 and 87 of this issue^{1,2}, we may hope for the emergence of a better verified picture, though no doubt still a complex one.

One of the observations on meiotic conversion that must be explained by hypotheses is that mutational sites often show a gradient of conversion frequency from one end of the gene to the other. This phenomenon, termed conversion polarity, has been taken to suggest that there is a conversion initiation site at one end of the gene and that the probability of a given mutational difference (genetic marker) being involved in conversion decreases with increasing distance from that site.

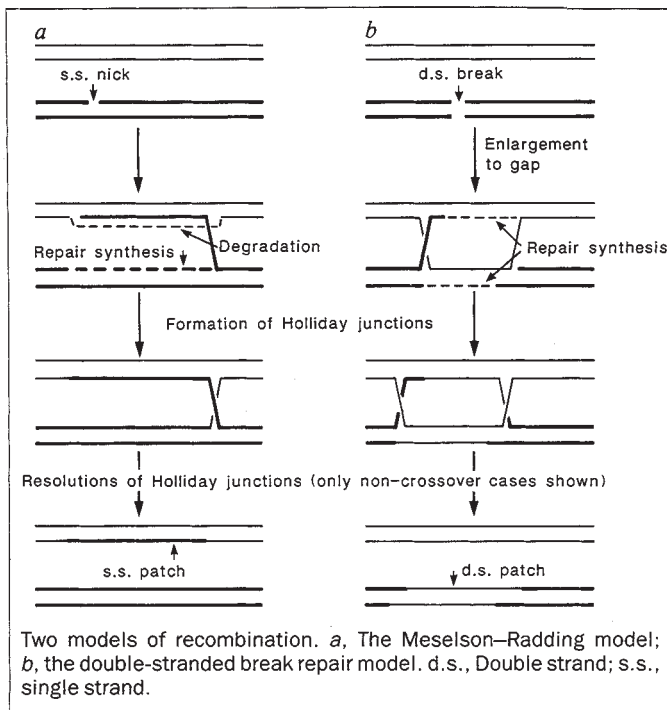
There are two radically different general hypotheses or models of recombination (see figure). Both involve the formation of single-strand exchanges (Holliday junctions) between homologous double-stranded DNA sequences, but they differ in the postulated origin of these junctions. Meselson and Radding's model³, which is an adaptation of Holliday's original concept, postulates the assimilation of single-stranded DNA from one duplex into the homologue, displacing the corresponding resident strand starting at a more-or-less fixed point. It is supposed that the displaced strand is destroyed and the single-strand gap in the donor duplex filled by new synthesis. If there is a mutational difference leading to non-complementarity between the invading strand and its new partner, mis-match correction may either restore the recipient to its original sequence or convert it to the sequence of the donor. If no correction

takes place, just one of the two products of the next round of DNA replication will show the conversion (post-meiotic segregation). In this model, the probability of conversion depends on the probability that the genetic marker is included within hybrid DNA which is pictured as extending for a variable distance from the initiation site; hence the closer the marker to the initiation site, the greater the chance of its being involved in conversion. In the original model⁴, the process of single-strand transfer was supposed to be initiated by single-strand nicking in the donor DNA, but another possibility is that it

be the precursor of at least those conversion events associated with post-meiotic segregation, and is now generally accepted as likely to be the immediate precursor of many conversion events. But the Szostak model can also accommodate heteroduplex at the margins of double-strand repair patches. Both invoke a site of preferential cleavage for the initiation of recombination, with the implication that deletion of this site should reduce recombination frequency. But whereas the double-strand gapped duplex in the double-stranded break repair model is necessarily the recipient of genetic information, the single-strand-nicked duplex in the original Meselson–Radding model is the donor. However, in Radding's alternative version of the Meselson–Radding model⁴, it is the single-strand gap in the recipient duplex that initiates conversion.

Both models suggest that the events leading to conversion are also responsible for crossing over, through one kind of resolution of the Holliday junctions. Thus both would give the putative initiation sites an essential role in all recombination, including exchanges of chromosome segments as well as conversion.

In the first of the two papers from Szostak's laboratory in this issue¹, Nicolas *et al.* present the first evidence at the molecular level concerning the mechanism of conversion polarity. The object of study is the yeast *Saccharomyces cerevisiae* ARG4 gene, in which conversion polarity was demonstrated many years ago by classical genetic methods⁵. First, Nicolas *et al.* introduced various kinds of mutations *in vitro* into defined positions in the cloned gene and, through the now



might be provoked by the appearance of a single-strand gap in the recipient.

The alternative hypothesis of Szostak *et al.*², termed double-stranded break repair, postulates that recombination is initiated by a double-stranded cut that may become a gap by erosion of the cut ends. According to this model, which is strongly supported by experiments on the recombinational behaviour of yeast plasmids containing a double-stranded gap, conversion is the result of repair of the gap by copying from both strands of the undamaged homologue. On this view, polarity could result, at least in part, from variable extents of gapping, with the chance of a marker being included in the gap decreasing with increasing distance from the site of gap initiation.

The two models are not very dissimilar in several of their predictions. The Meselson–Radding model most obviously provides for the heteroduplex which must

routine tricks of yeast molecular genetics, substituted these artificial mutations into the yeast genome. Genetic analysis of crosses between each of these defined mutants and the wild type show a gradient of conversion frequency essentially identical to that previously demonstrated for *arg4* mutations *in vivo*.

In the crucial second phase of their investigation, Nicolas *et al.* made deletions of various segments in and adjacent to the upstream (high-conversion) end of the gene, and tested them for their effects on conversion frequencies of sites at different distances downstream. Deletions covering the region from 30 to 300 bases upstream of the transcription initiation site, when they are present in both parents of a cross, abolish the polarity and reduce the conversion frequency at all sites — not to zero but to a level similar to that which is normally found at the low end of the conversion gradient.

The authors suppose that the main *ARG4* recombination initiation site lies in the -300 to -30 segment, which also happens to include the *ARG4* promoter. When they brought a marker normally at the low end of the conversion gradient close to the promoter by deleting most of the intervening sequence, its conversion frequency was increased to the level characteristic of the high end of the gradient. They attribute the residual conversion activity following the deletion of the putative conversion initiation region to recombination initiated at other, probably more distant sites. When the putative initiation region is deleted in one parent but not in the other, it is the non-deleted partner that is the preferred recipient of information in conversion; there is a large excess of 3:1 over 1:3 ratios of deleted to wild-type products (conversion disparity).

This is, of course, what the double-stranded break repair model would predict. In a substantial proportion of such 3:1 tetrads, a marker several hundred base pairs downstream of the deletion is included in the conversion event, as if the gap extended for at least that distance from the initiation site. Deletions not covering the promoter/initiation region show no conversion disparity, or a disparity not held to be statistically significant (a somewhat questionable conclusion in one case).

In the second paper in this issue from Szostak's group, Sun *et al.*² report on the fate, in a diploid culture synchronized in meiosis, of replicating plasmids carrying cloned chromosomal DNA, including the *ARG4* gene, with the putative initiation region either present or deleted. Restriction fragment analysis shows that, when this region is present, three double-strand breaks appear at different sites in the cloned DNA; these are all in, or close to, known promoter regions and one of them (site 2) is in a position consistent with that of the putative *ARG4* recombination initiation site. Each of the breaks affects only a small proportion of the plasmid molecules, perhaps not too different from the observed frequency of gene conversion in *ARG4*. They appear early in meiosis and disappear as meiosis is completed. Plasmids with the putative *ARG4* initiation region deleted are still cut at sites 1 and 3, but not at site 2. These observations provide strong evidence that the recombination initiation site is readily recombable in meiosis and suggests

that a double-strand cut may be the initiating event.

Another observation of considerable interest³ is that the fragments are substantially shortened by single-strand exonuclease, implying that they have long single-stranded tails. Assuming that these tails are not artefacts of cell disruption, this observation suggests that the gap-

repair process involves the formation of longer tracts of heteroduplex than had previously seemed likely. This would give the double-stranded break repair model still more versatility in explaining the genetic phenomena. □

J.R.S. Fincham and P. Oliver are in the Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK.

METEOROLOGY

Cold waters and hot summers

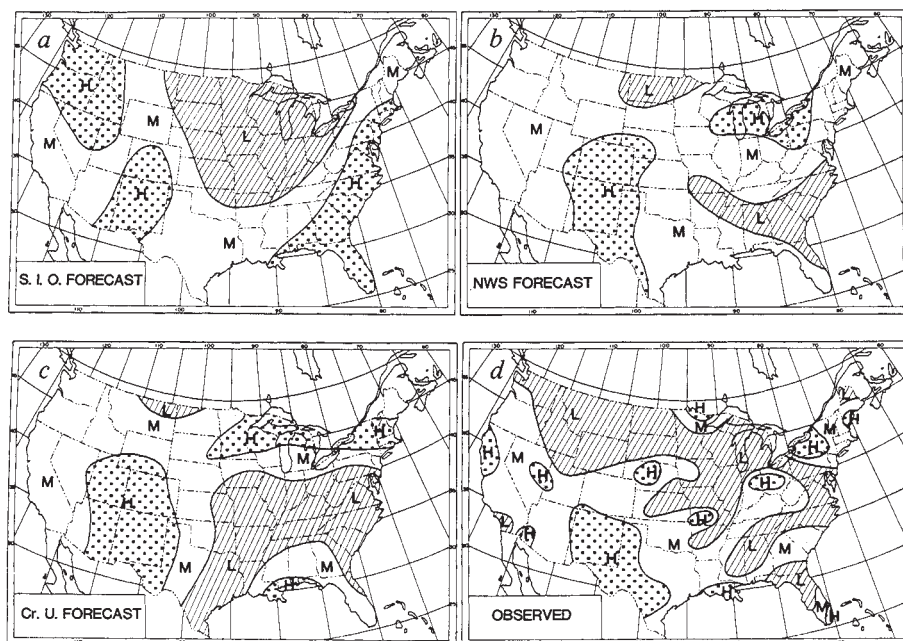
Jerome Namias

THE great US drought of the summer of 1988 was linked to sea-surface-temperature changes, according to Palmer and Branković on page 54 of this issue¹. Their work extends the recent studies by Trenberth *et al.*² who attributed the drought mainly to the arrival of exceptionally cold water in the tropical Pacific (La Niña) in the aftermath of the warm El Niño event of 1987. The novelty of these two contributions is not that abnormal sea-surface-temperature patterns are necessary links in the climate system during the droughts in the United States, but that a surface-temperature anomaly in a specific region is the main culprit.

Empirical studies of many great droughts of this century have demonstrated that summer drought over the United States is part of a larger picture in which three upper-level high-pressure cells — over the North Pacific, the North Atlantic and the United States — interact in a mutually supporting fashion. The term teleconnections has been applied to

such interactions — a phenomenon first explored by Sir Gilbert Walker³.

Palmer and Branković¹, like Trenberth *et al.*², have used state-of-the-art numerical models to make their case. Trenberth *et al.* showed that the three-cell flow pattern at a height corresponding to 300 millibars (in the upper troposphere) is reasonably well simulated if the model is initialized with the abnormally cold sea surface temperature that existed in the tropical North Pacific during the drought. Palmer and Branković compare model global simulations and 30-day weather predictions made starting with the conditions in May 1987 (when no comparable drought existed) with those initialized with the conditions in May 1988. Both papers indicate that sea-surface-temperature boundary conditions were necessary links in the establishment of drought in North America and conclude that the primary cause lay in the tropics where La Niña occurred in 1988. (La Niña is a period characterized by the presence of



Forecasts of (a-c) precipitation classes for summer 1988; a, by myself and D. Cayan; b, the Long Range Prediction Group of the US National Weather Service; and c, Douglas of Creighton University. d, The observed patterns. Classes are terciles computed from 30 summers. L, light; M, moderate; and H, heavy. The models predicted the drought, but not its severity.

1. Nicolas, A., Treco, D., Schultes, N.P. & Szostak, J.W. *Nature* **338**, 35-39 (1989).
2. Sun, H., Treco, D., Schultes, N.P. & Szostak, J.W. *Nature* **338**, 87-90 (1989).
3. Meselson, M. & Radding, C.M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 356-361 (1985).
4. Radding, C.M. *et al.* *Cold Spring Harb. Symp. quant. Biol.* **47**, 821-828 (1982).
5. Szostak, J.W. *et al.* *Cell* **33**, 26-35 (1983).
6. Fogel, E. *et al.* in *The Molecular Biology of the Yeast Saccharomyces* (eds Strathern, J. *et al.*) 289-339 (Cold Spring Harbor, New York, 1981).

anomalously cold surface water in the North Pacific Ocean, and frequently follows El Niño, a period of anomalously warm waters. The cause of these events is unknown, but they have been associated with many types of climate change.)

I believe the conclusion these groups draw is too strong, and argue elsewhere¹ that La Niña is only one of several factors, and by no means the strongest one, associated with the drought through the three-cell structure. Indeed, Trenberth *et al.* and Palmer and Branković acknowledge that this localization of the primary cause is far from conclusively demonstrated by the model results, even though the effect of hemisphere-wide sea-surface-temperature anomalies is strongly suggested. The anomaly patterns are closely related to atmospheric flow patterns, and the cold La Niña water and the warm pool of water just to its north could be either causes of or responses to remote events. Also, Palmer and Branković's data are at best suggestive: there is a disturbing spatial offset between the differences in 300-millibar height patterns and temperature and precipitation patterns over North America and adjacent areas predicted, and those observed.

As both sets of authors point out, the drought of 1988 and its causes could be unique. Many major North American droughts of this century have not been associated with La-Niña-type conditions, although the three-cell pattern has occurred. Dry soil in spring in the mid-west, which generates a high-pressure cell², is conducive to drought (as Trenberth *et al.* acknowledge) and was a factor in 1988. And warm water in either the North Atlantic or North Pacific would also generate high-pressure, anticyclonic cells. It is possible that in the present case, the La Niña event was so strong that it simply eclipsed the other air-sea-land interactions. Thus as is found in other similar studies, cause and effect are hard to distinguish.

The influence of the tropics on atmospheric general circulation, as highlighted by these new studies, is clearly important. Although the drought last year was foreseen in empirical models its severity was underestimated (see figure). Future predictions could benefit from our improved understanding. One thing, however, is absolutely clear: the drought was a consequence of normal atmospheric variability, and has no connection whatever with the greenhouse effect. □

Jerome Namias is at the Scripps Institute of Oceanography, La Jolla, California 92093, USA.

Detecting density dependence in imaginary worlds

Robert M. May

CHAOS — in the form of simple and fully deterministic systems that exhibit very complicated and effectively unpredictable dynamics — is much in the news, partly, I think, as a result of Gleick's popular book¹. Although the practical implications of deterministic chaos were first recognized by Lorenz² in a meteorological context and by Yorke, Oster, myself and others³⁻⁵ in dynamics of animal populations, most of the recent excitement centres on applications to chemical, physical, electrical and physiological systems, where cascades of period doubling and other related phenomena can be observed in a precise and satisfying way. Among the many reasons for the shift in emphasis is the fact that the magnitudes of real populations of plants and animals are rarely governed by simple nonlinear equations that could provide crisp illustrations of chaotic behaviour. Rather, such populations are typically affected by interactions with many others and by unpredictable environmental fluctuations. The result is a confused picture, with no general agreement about the implications of chaotic phenomena for biological populations^{6,7}. A recent paper by Mountford⁸ sheds some light on these questions.

One of the uncertainties concerns the extent to which strong nonlinearities and the possibility of chaotic dynamics in natural populations might, in certain circumstances, invalidate conventional methods of analysing data. In the past, the implicit assumption has been that density-dependent effects would regulate a population to some constant value, or at worst to a stable cycle, were it not for unpredictable fluctuations in environmental and demographic parameters. The task thus seemed to be one of extracting a regulatory 'signal' from obscuring 'noise', and conventional techniques (such as *k*-factor analysis, which essentially seeks patterns in the relation between population densities in successive generations) are implicitly based on this. But what if the observed fluctuations in population densities derived partly from real 'noise' and partly from deterministic chaos (apparently random fluctuations produced by a simple density-dependent relation)?

Hassell⁹ has attacked this question by examining the extent to which standard techniques of data analysis can uncover the rules obeyed by imaginary populations of his own creation. This general idea is useful in other areas of ecology, for example Colwell and Winkler's exploration of the extent to which prevailing methods could uncover the biogeographical rules

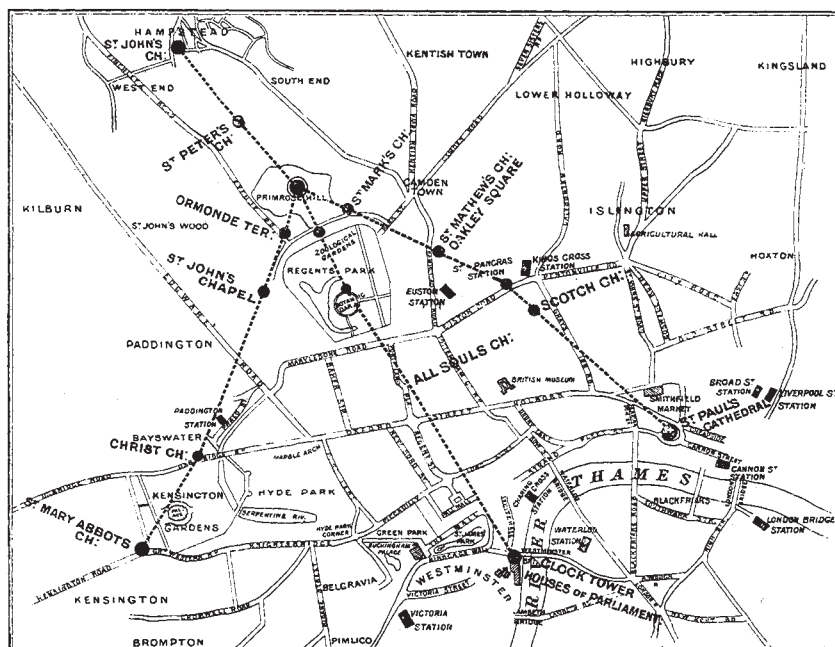
whereby their computer program colonized simulated archipelagos of islands with simulated species^{10,11}. Hassell's 'pseudo-data' were generated by a program that embodied ideas about how some insect populations may behave: in each generation adults were distributed (according to a statistical recipe) among many different patches; in each patch, the adults laid eggs; larvae then emerged, and their probability of survival to adulthood depended, in a specified way, on the density in that patch; in this way, the next generation of adults was produced.

Hassell also superimposed various kinds of stochastic fluctuations (in the number of eggs laid by individual adults, for example, or in the patterns of distribution of adults among patches) on his system. If the density-dependent effects acting on larval survival were sufficiently severe, then Hassell's pseudo-data showed irregular fluctuations that were a mixture of the genuine noise he introduced and oscillatory or chaotic fluctuations produced by density-dependent effects. Even when density dependence regulated the population to a constant value in the absence of environmental noise, the stochasticities introduced by Hassell could lead to fluctuations in parameters influencing the severity of density-dependent effects, which in turn could cause the sporadic manifestation of chaotic dynamics.

Applying *k*-factor analysis to these pseudo-data, Hassell could sometimes detect the underlying density-dependent mechanisms that ultimately regulated his populations, and sometimes could not. That is, in this system where the overall density in any one generation depended on a patchy distribution, where density-dependent effects acted differently at different densities in different patches, and where superimposed noise of various kinds made chaos possible in some patches, the regulatory 'signals' were masked by some kinds of noise, but not by others. Hassell *et al.*¹² have now applied these insights to an analysis of about 20 years of data on the abundance of whitefly populations on leaves on viburnum bushes at the Imperial College Field Station at Silwood Park, and seem to have detected density-dependent effects that were not revealed by conventional analysis of the data for overall average whitefly populations in successive generations. But these preliminary studies leave open many questions about which kinds of environmental noise do and do not mask density-dependent signals in Hassell's simulations,

1. Palmer, T.N. & Branković, Č. *Nature* **338**, 54–57 (1989).
2. Trenberth, K.E., Branstator, G.W. & Arkin, P.A. *Science* **242**, 1640–1645 (1988).
3. Walker, G.R. & Bliss, E. *World Weather V. Mem. R. met. Soc.* **4**, 54–84 (1932).
4. Namias, J. *J. Clim.* (in the press).
5. Shulka, J. & Mintz, Y. *Science* **215**, 1498–1500 (1982).

100 years ago THE DARKNESS OF LONDON AIR



The constitution of London fogs has been carefully gone into by several well-known men of science; and the results obtained are of very great interest, as they prove, amongst other things, that during the winter London air has an unusually large amount of carbonic acid in it. Various observations were taken in London during the winter of 1887-88. During the five months selected, Christ Church, Lancaster Gate, and St. Mary Abbot's Church, Kensington, on the south-west line; the Clock Tower, Houses of Parliament, on the south line, and the Scotch Church, Regent's Square; and St. Paul's Cathedral, on the south-east line, were *never once* seen. When it is known that on any fine day during the late spring, summer, and early autumn, you can see right across London, on any one of the selected lines, it will be easy to realize how thick the air over London is during the winter. From *Nature* 39, 442; 7 March 1889.

much less in the real world.

In his new paper⁸, Mountford has extended this discussion. Hassell assumed that the competitive effects determining larval survival are 'scramble': when resources are abundant, all do well, and when resources are sparse, all do badly. The result can be a highly nonlinear, 'boom-and-bust' relation, which can fairly easily produce oscillatory or chaotic dynamics. Mountford, on the other hand, bases all his simulations on density-dependent relations corresponding to 'contest' competition, which assumes resources are distributed in a hierarchical fashion, so that a few individuals do well even in hard times. These smooth relations chosen by Mountford to characterize larval competition can never generate oscillatory, much less chaotic, dynamics. With his pseudo-data generated in this way, Mountford finds that conventional methods can easily detect the density-dependent effects, with spatial heterogeneity and environmental noise presenting no difficulties.

The problems discovered by Hassell seem to me to depend largely on the interplay between what might be called 'density-dependent noise' (generated by nonlinear relations that, in some patches in a fluctuating environment, can be severe enough to produce chaos) and 'density-

independent noise' (generated directly by environmental stochasticity), in a patchy world¹³. So I am not surprised that these problems do not arise in systems such as Mountford's, where the nonlinearities are too weak to produce oscillations or chaos. It is perhaps unfortunate that Mountford does not mention this qualitative difference between his chosen density-dependent functions and the boom-and-bust ones of Hassell; the differences between Hassell's and Mountford's simulations seem puzzling if this underlying difference is not appreciated.

This being said, Mountford's study is interesting in several ways. First, it is useful to have an explicit indication that patchiness and stochastic effects, by themselves, are insufficient to upset traditional methods of data analysis. Severe nonlinearities in density-dependent effects thus seem to be an essential ingredient in such problems as may exist. Second, Mountford shows that (in the absence of strong nonlinearities) spatial heterogeneity can actually enhance our ability to detect density dependence: "if there are a number of sets of population density series all with the same mean and the same variability but with differing degrees of spatial heterogeneity, then the detection of density dependence improves with increasing spatial heterogeneity". Third,

the pronounced differences between Hassell's studies (based on 'scramble' competition, with its potential for strong nonlinearities and chaotic dynamics) and Mountford's (based on 'contest' competition) highlight the often-neglected way in which the behaviour of individuals influences the dynamics of populations. Indeed, Lomnicki's recent book¹⁴ is devoted to showing explicitly how some kinds of foraging (or other) behaviour result in 'scrambling' competitive relations and thus in highly volatile dynamics for the population, while different kinds of individual behaviour result in 'contest' competition and thus in tame dynamics. The comparison between Hassell's and Mountford's papers provides another perspective on Lomnicki's basic theme.

There seems to me to be much scope for further studies of ways in which spatial heterogeneity, environmental unpredictability, and nonlinear interactions within and between populations can swirl together to confound empirical studies aimed at understanding what prevents the long-term average density of a population from increasing indefinitely (or decreasing to zero in a time short compared with average extinction times). Using computers to generate pseudo-data for imaginary worlds whose rules are known, and then testing conventional methods of data analysis for their efficiency in revealing these known rules, seems to me to be a useful approach. We must all, of course, join Mountford in agreeing with Dempster and Pollard¹⁵ that "the best hope of unravelling the roles of different factors in the population dynamics of animals, still rests in analysis of long-term, life-table data". The worry remains that, until we have a surer grasp of possible complications in the analysis, we cannot be certain we have gathered the appropriate data, no matter how long and carefully we have toiled in the field. □

Robert M. May is a Royal Society Research Professor at Imperial College London and at the Department of Zoology, University of Oxford, Oxford OX1 3PS, UK.

- Gleick, J. *Chaos: Making a New Science* (Viking, New York, 1987).
- Lorenz, E.N. *J. Atmos. Sci.* **20**, 130-141 (1963).
- Li, T.Y. & Yorke, J.A. *Amer. Math. Mth.* **82**, 732-733 (1975).
- May, R.M. *Science* **186**, 645-647 (1974).
- May, R.M. & Oster, G.F. *Am. Nat.* **110**, 573-599 (1976).
- Schaffer, W.M. & Kot, M. *Trends Ecol. Evol.* **1**, 58-63 (1986).
- Berryman, A.A. & Millstein, J.A. *Trends Ecol. Evol.* **4**, 26-28 (1989).
- Mountford, M.D. *J. anim. Ecol.* **57**, 845-858 (1988).
- Hassell, M.P. *J. anim. Ecol.* **56**, 705-713 (1987).
- Colwell, R.K. & Winkler, D.W. in *Ecological Communities: Conceptual Issues and the Evidence* (eds Strong, D.R., Simberloff, D., Abele, L.G. & Thistle, A.B.) 344-359 (Princeton University Press, 1984).
- Harvey, P.H. & May, R.M. *Nature* **314**, 228-229 (1985).
- Hassell, M.P., Southwood, T.R.E. & Reader, P.M. *J. anim. Ecol.* **56**, 283-300 (1987).
- May, R.M. *Proc. R. Soc. B* **228**, 241-266 (1986).
- Lomnicki, A. *Population Ecology of Individuals* (Princeton University Press, 1988).
- Dempster, J.P. & Pollard, E. *Oikos* **46**, 4413 (1986).

Big whorls do have little whorls

Ian Stewart

IN 1922 Lewis Fry Richardson wrote *Weather Prediction by Numerical Process*, and it won him election to the Royal Society. For 20 years it was considered disreputable, because of problems with the numerical methods proposed, but by 1955 it had become a classic. It contains a much-quoted parody of Jonathan Swift:

Big whorls have little whorls,
Which feed on their velocity;
And little whorls have lesser whorls,
And so on to viscosity
(in the molecular sense).

This is the famous Richardson cascade, a descriptive theory of turbulence. Although the mechanism is intuitively attractive, the experimental evidence in its favour has until now been largely indirect. Elsewhere in this issue (*Nature* 338, 51–53; 1989) Argoul *et al.* provide the first experimental geometric evidence for the cascade, and in particular for Benoit Mandelbrot's suggestion that the process is fractal.

If a drop of ink is placed in a jar of water, it creates a vortex ring as it falls. The ring acquires corners, from which smaller vortices break off. The process repeats to form a tree-like structure of

ever-smaller vortices. Each stage is geometrically similar to the previous one, but on a smaller scale — a structure characteristic of fractals. Richardson's idea is that something akin to this proliferating tree of vortices occurs in turbulent fluids. Large vortices break up into smaller vortices — presumably because they are unstable — which themselves break up into still smaller vortices.

The total energy is reduced by viscous friction, but the energy loss is slow to begin with. It becomes much more rapid as the size of vortices decreases, hence Richardson's parenthetical remark. Thus the energy of the fluid cascades into ever-smaller vortices until it is damped out by viscosity.

This suggestion is not Richardson's alone. Similar ideas go back to L. Prandtl, G. I. Taylor, and the statistical theory of turbulence of A. N. Kolmogorov (see A. A. Townsend, *The Structure of Turbulent Shear Flow*; Cambridge University Press, 1976). But how accurate a picture of turbulent flow is it? On physical grounds the idea is fairly plausible; but it has proved extremely difficult to make experimental measurements that are suf-

ficiently precise to justify the model in detail. There is also the difficult problem of extracting the intricacies of vortex dynamics from measurements of a few variables.

Argoul *et al.* overcome many of these difficulties by using a new analytical technique, the wavelet transform. The traditional Fourier transform represents a waveform in terms of a series of sine curves. Instead of sines, the new method represents the waveform in terms of localized solitary waves, more appropriate for nonlinear dynamics. The wavelet transform includes a scaling factor, and can therefore analyse fractal properties. In a sense, it renders visible the construction procedure for the fractal: in this case, the way that a vortex splits into smaller vortices. Successive splittings show up as successive branchings of the graphical representation of the transform.

Argoul *et al.* observe a tree-like branching pattern in high-velocity turbulence, similar to the construction procedure for one of the simplest fractals: the Cantor set. This set is universally misattributed: it seems to have been invented by Henry Smith in 1875, but Cantor drew attention to it in 1883 — citing Smith's work. It is obtained by repeatedly deleting the middle thirds of intervals, and thus has left-right symmetry.

As is made clear in the graphical

Guanylate cyclases send out new signals

GUANYLATE cyclase has long had to play second fiddle to adenylate cyclase. Both enzymes were discovered in the wake of the identification, 30 years ago, of cyclic AMP as the intracellular mediator of the activation of glycogen metabolism by extracellular adrenaline. Adenylate cyclase and its product, cyclic AMP, have since been established as a much-used system of generating intracellular signals in response to the docking of hormones onto their cell-surface receptors. By comparison, the equivalent system of guanylate cyclase and cyclic GMP is under-employed and, perhaps consequently, often overlooked. But two separate strands of research are focusing renewed attention on the system.

Unlike adenylate cyclase, which is invariably part of the cell membrane, guanylate cyclase can be either membrane-bound or in soluble form in the cell cytoplasm. The membrane form has been intensively studied since the discovery, in the past few years, of the family of atrial natriuretic peptides (ANPs), which use cyclic GMP to mediate their signals that result in the excretion of sodium (natriuresis) and water, and in vasodilation. There has been mounting evidence that the receptor for ANP is, itself, a membrane form of guanylate cyclase. On page 78 of this issue¹, Michael Chinkers *et al.* provide the ultimate evidence: the complementary DNA

for the enzyme from rat brain, when expressed in cells, has both guanylate cyclase and ANP-binding properties.

The deduced sequence of the enzyme is consistent with the obvious notion that the ANP-binding portion is on the exterior surface of the membrane and the guanylate cyclase on the interior. The sequence of the cyclase portion has considerable similarity with a membrane form of guanylate cyclase that mediates responses of sea-urchin sperm to peptides released by eggs. This observation is hardly surprising, as a complementary DNA for the sperm enzyme² was used as the probe to isolate the rat brain complementary DNA.

More interestingly, the sequence of the guanylate cyclase portion of the rat brain membrane enzyme has considerable similarity with the only sequence available for a soluble form of guanylate cyclase³, although Chinkers *et al.* conclude that the latter must be the regulatory, rather than the catalytic, subunit of the soluble enzyme.

One striking feature of the soluble enzyme not shared by the membrane form is that it contains a haem group. Some evidence has suggested that it is by interacting with the haem that endothelium-derived relaxing factor, now known from the work of Salvador Moncada and his colleagues⁴ to be nitric oxide, activates the soluble guanylate cyclase of smooth

muscle, thereby generating the cyclic GMP that mediates muscular relaxation.

That observation created a flurry of interest and a subsequent broadening of its relevance. It is now known that L-arginine is the endogenous precursor of nitric oxide in endothelial cells^{5,6} and in macrophages⁷, where it is clearly an intermediate in the synthesis of nitrite, but will now also have to be investigated as a possible intracellular signal of macrophage activation by means of guanylate cyclase stimulation. Evidence has also emerged that nitric oxide (not positively identified) mediates the increase of cyclic GMP that follows the interaction of external glutamate with its receptors (of the NMDA type) on the surface of cerebellar cells⁸.

Whether or not guanylate cyclases are destined always to play second fiddle to adenylate cyclases in the events that orchestrate cellular responses to external stimuli, they seem certain to make themselves heard in the coming months.

Peter Newmark

Peter Newmark is Deputy Editor of *Nature*.

1. Chinkers, M. *et al.* *Nature* **338**, 78–83 (1989).
2. Singh, S. *et al.* *Nature* **334**, 708–712 (1988).
3. Koesling, D. *et al.* *FEBS Lett.* **239**, 29–34 (1988).
4. Palmer, R.M., Ferrige, A.G. & Moncada, S. *Nature* **327**, 524–526 (1987).
5. Palmer, R.M. *et al.* *Nature* **333**, 664–666 (1988).
6. Sakuma, I. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 8664–8667 (1988).
7. Marletta, M.A. *et al.* *Biochemistry* **27**, 8706–8711 (1988).
8. Garthwaite, J., Charles, S.L. & Chess-Williams, R. *Nature* **336**, 385–388 (1988).

GEOCHEMISTRY

Tales of a lost magma ocean

Laura Garwin

representation of their results, the Richardson cascade observed by Argoul *et al.* is asymmetric. In other words, during the cascade, each vortex splits into smaller vortices of different sizes. Moreover, the scaling factor does not remain constant throughout the process, so the particular Richardson cascade that they observe is not precisely self-similar.

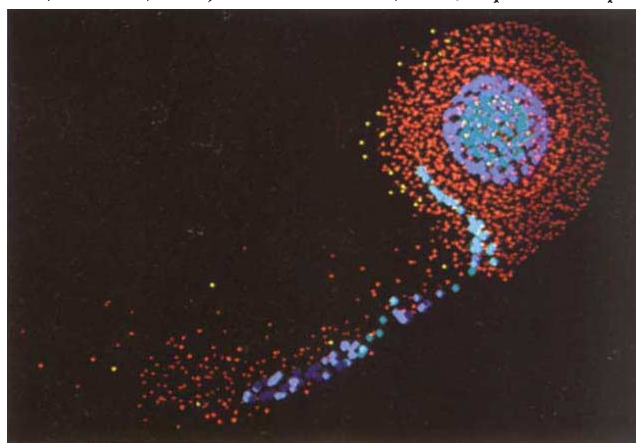
Recall that every fractal has associated to it a number, its Hausdorff dimension, which measures how irregular it is. Typically this dimension is not a whole number. Introductions to fractal geometry tend to emphasize self-similarity — that apart from the scaling factor the object looks the same at different levels of detail — and define the dimension in terms of constant scaling. But this property is not a prerequisite for being a fractal. In particular the dimension can be defined for fractals that do not have uniform scaling features. On one popular definition a fractal is just a set whose Hausdorff dimension differs from its topological dimension; less technically, it must possess detailed structure on all scales. The simplest way to achieve this is to have the same structure on all scales — that is, to be self-similar — but there exist many fractals whose structure varies from one scale to another. To emphasize this feature, fractals that are not self-similar are often referred to as multi-fractals.

For a rough analogy, modify the Cantor-Smith construction. Instead of always removing the same proportion of the interval, first delete the middle third, then the middle quarters of the remaining subintervals, then middle fifths, sixths, sevenths and so on. Because the deleted proportions vary, there is no true self-similarity, so the modified Cantor-Smith set is a multi-fractal. And to introduce asymmetry, delete intervals that are not centrally placed.

Thus there now exists experimental evidence for the fractal nature of turbulence, clarifying details of the fractal structure, and using this to show that the splitting of vortices is asymmetric and only approximately self-similar. These are features that must therefore be incorporated into any realistic model of the Richardson cascade, or deduced from any more detailed theory of turbulence. These are important results. However, the theory that they confirm is purely phenomenological. The Richardson cascade describes features of turbulent flows — but it does not derive them from the equations of fluid motion. Although fractals are typically associated with chaos in non-linear dynamics — apparently random motion in deterministic systems — these new results bear only indirectly on the relation between turbulence and chaos. □

Ian Stewart is at the Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

CURRENT models for the formation of the terrestrial planets from the dust and gas of the solar nebula postulate a growing population of planetesimals, which in the final stages of accretion may have included about 100 Moon-sized bodies, ranging up to a few the size of Mars. Collisions between these proto-planets have been invoked to explain anomalies in the present-day Solar System, such as the retrograde rotation of Venus, the exceptional tilt of Uranus and the very high density of Mercury (see the recent News and Views article by G. Stewart, *Nature* **335**, 496–497; 1988). Closer to home, the



Computer simulation of the impact of a Mars-sized body with the Earth, showing the configuration about 80 min after initial contact. Rock particles are in shades of red and yellow, iron particles in blue and green. Lighter shades denote higher internal energy (a measure of the degree of heating or compression following impact).

orbital characteristics of the Earth-Moon system can also be explained by a 'giant impact' — in this case, a glancing collision between a Mars-sized body and the proto-Earth, with the Moon forming from the debris of the impact. Elsewhere in this issue (*Nature* **338**, 29–34; 1989), H.E. Newsom and S.R. Taylor outline the consequences of such an event for the bulk chemistry of the Earth and Moon, and conclude that the hypothesis is cosmologically plausible. But given the enormous energy involved in such a collision, might we not expect to see some record of it preserved in the Earth's present structure? A recent conference* brought impact modellers together with those willing to speculate about conditions on the earliest Earth, to examine the evidence for or against the giant impact hypothesis.

The total energy available from the impact of a Mars-sized body on the proto-Earth is of the order of 5×10^{31} joules — enough to raise the temperature of the entire Earth by about 10,000 K. But it is

not obvious that this energy will be transferred efficiently to the Earth, or distributed uniformly within it. Jay Melosh (University of Arizona) addressed this question by considering the heating caused by shock waves propagating away from the point of impact. His numerical calculations give temperature rises of 3,000–4,000 K throughout the Earth's mantle for an impactor whose initial velocity v (at large separation) is 7.8 km s⁻¹; for $v=0$, at least half the mantle (the hemisphere adjacent to the impact) is heated this strongly. Given a mantle liquidus temperature of about 3,000 K,

even an initially cold Earth would have a completely or substantially molten mantle after the impact.

Independent calculations presented by Al Cameron (Harvard-Smithsonian Center for Astrophysics) yield similar results, and also indicate that the Earth's surface temperature would remain very high for thousands of years, as the rock-vapour atmosphere created by the most intensely heated material (see figure) accreted onto the Earth's surface. The mantle could

remain molten for even longer times, if enough volatiles re-accreted to the Earth to form a dense atmosphere.

The feldspar-rich highland rocks of the Moon are thought to have crystallized from a 'magma ocean' more than 100 km deep; if the Earth had its own magma ocean, where are its crystallization products? Unlike the Moon, the Earth is tectonically active, so one would not necessarily expect the products themselves to have survived. But a central tenet of modern geochemistry is that whenever crystals separate from a melt, either during melting or crystallization, a characteristic imprint is left on the chemistry of the melt and crystals. This arises because some elements (the incompatible, or trace, elements) fractionate preferentially into the melt, whereas others prefer the solid phases. A melt can crystallize to a solid of exactly the same composition only if the crystals remain in continuous contact with the melt, so that the chemical system is closed, or if all elements show no preference between crystals and melt (all partition coefficients are unity).

Given the bulk composition of the

*Conference on the Origin of the Earth, Berkeley, California, 1–3 December 1988.

Earth's mantle, the depth (pressure) and temperature range of the molten layer, and partition coefficients for diagnostic trace elements between the melt and the minerals that would crystallize from it at those pressures and temperatures, it should be straightforward to calculate the effects of a crystallizing magma ocean on the trace-element contents of the minerals involved. This is the approach taken by Ted Ringwood (Australian National University), who has measured high-pressure partition coefficients for various elements, and used them to model the crystallization of magnesium-perovskite from a pyrolite melt. He concludes that a primitive crust would have formed, which was strongly enriched in some elements and relatively depleted in others.

As a result, certain pairs of elements (for example, lutetium and hafnium) would be strongly fractionated relative to one another; yet Ringwood noted that the oldest minerals known (4.2×10^9 -year-old zircons) have chondritic (undifferentiated) Lu/Hf ratios. He concluded that there could not have been a terrestrial magma ocean, unless the Earth was re-homogenized by 4.2×10^9 years ago, removing all trace of its existence.

This seeming contradiction between Ringwood's experiments and the existence of a substantially molten early Earth was the subject of much controversy, as other speakers invoked independent evidence for a global magma ocean, or questioned the assumptions and experimental data behind Ringwood's model. Carl Agee (University of Bayreuth) addressed the problem of obtaining a terrestrial upper mantle of peridotitic (olivine-rich) composition from a primordial Earth with the more silica-rich composition of the CI chondrites (primitive meteorites which are thought to represent the composition of the solar nebula). A simple mass-balance argument shows that upper-mantle peridotite can be formed from a CI composition by subtraction of Mg-perovskite and addition of olivine — a combination that can be achieved by flotation of olivine and settling of perovskite crystals in a mantle molten to at least 1,000-km depth. Agee suggested that the observed partition coefficients for elements such as Al, Ti, Hf, Lu, Zr, Sc and Sm were consistent with the 30 per cent fractionation of perovskite required by his mass balance, but this point was hotly contested by Ringwood. Mike Drake (University of Arizona) considered that the addition of 30 per cent olivine would result in substantially non-chondritic ratios of Sc/Sm, Ni/Co and Ir/Au, which are not observed.

There was much debate concerning the reliability of partition coefficients measured at very high pressure — the experiments are difficult to perform, and perhaps even more difficult to interpret —

and about their application to chemical systems more complicated than those of the experiments. But Al Hofmann (Max Planck Institute, Mainz) made the general point that the partition coefficients are bound to deviate from unity, "whether you measure them or not". What must be explained is the general lack of fractionation of trace elements (preservation of chondritic ratios), in the face of a major-element composition so different from chondritic that significant differentiation must have occurred.

In an illustration of the principle that "for each geochemist there exists an equal and opposite geochemist", Herbert Palme (Max Planck Institute, Mainz) pointed out that we do not know that the Earth started out with a CI chondrite composition. He argued that, if the upper mantle is representative of the entire mantle, the Earth's bulk composition is closest to that of the CV chondrites, which have lost silica relative to the CIs, probably in the solar nebula. This proposal avoids the need for extensive differentiation of the Earth's mantle, although there may be problems in detail with CV chondrite as a parent material for the Earth.

In contrast, Don Anderson (California Institute of Technology) put his faith in "seismic chemistry" as a more reliable guide to the composition of the bulk mantle than samples of upper-mantle

peridotite. He cited density and seismic-velocity data supporting an FeO content of about 14 per cent for the lower mantle, as compared with values of 8–12 per cent for the upper mantle. Although this FeO content is high compared to chondrites, it is comparable to a recent measurement for the solar photosphere (H.H. Breneman & E.C. Stone *Astrophys. J.* **289**, L57; 1985), and to the values inferred for the Moon and Mars. So this line of argument forces us once again to explain the Mg-rich upper mantle by large-scale chemical differentiation, and to confront the paradox of unfractionated trace elements.

Returning to the question that prompted the conference, it seems that mantle geochemistry cannot (yet?) be used to support or contradict the giant impact hypothesis. As Anderson emphasized, Ringwood's experiments (or, more generally, partition coefficients substantially different from unity) do not rule out melting; they only rule out fractionation at mantle pressures. Thus, if crystals could be kept in contact with the melt throughout crystallization, for example by entrainment in a turbulently convecting flow (W. B. Tonks, University of Arizona), a global magma ocean could come and go without leaving any imprint on the chemistry of the mantle. □

Laura Garwin is Physical Sciences Editor of Nature.

POLYMERS

Surface segregation of blends

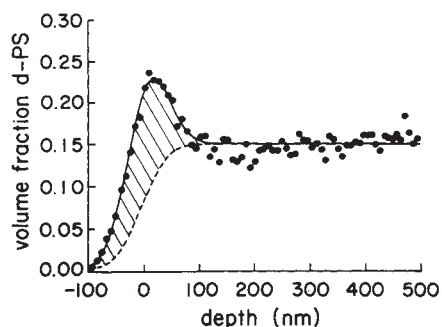
A. M. Donald

Ask physicists what length scale they associate with surface phenomena and they are likely to come up with a value of 1 nanometre or less. But for polymers, because of their large chain length, the distance associated with a surface 'monolayer' is much larger, and corresponding timescales much longer than for simple atomic or molecular systems. Thus polymer systems could be ideal for experimental tests of recent theoretical predictions concerning wetting and surface phase transitions in binary mixtures, which relate directly to the precise composition of the surface. Few experiments have yet been carried out to look at the surface composition of polymer blends (miscible two-component systems) but new results from E. J. Kramer and colleagues^{1,2} indicate that a very marginal difference between the surface energy of polystyrene and its deuterated form gives rise to a dramatic enrichment of the latter at the surface relative to the bulk composition. It is to be expected that even more striking effects could be seen with blends containing chemically distinct polymers.

Few polymer pairs are miscible. This can be explained in terms of the small

entropy of mixing relative to the total entropy: the number of configurations in a long chain system is already so large there is little to be gained by mixing in a second species. Thus mixing will occur only when there is a sufficiently favourable specific (enthalpic) attraction between the two components. A particularly simple blend system arises when the two components differ only by isotopic substitution, for example of all the hydrogen atoms (¹H) on the chain by deuterium (²H). Polystyrene blends become miscible above a critical temperature (176 °C for the molecular weights used in the study of Kramer and co-workers), and thus in the bulk there is a homogeneous distribution of the two components at temperatures greater than this.

But what happens in the vicinity of a surface of a miscible blend? If the two components have different surface energies, then a driving force exists to permit the component of lower surface energy to segregate preferentially at the surface — as is known to occur for binary alloys and liquids of small molecules. The technique that Kramer and co-workers have used, forward recoil spectrometry



Volume fraction of deuterated polystyrene (d-PS) blended with polystyrene as a function of depth, after annealing at 184 °C for 5 days (initial volume fraction 0.15). Shaded area, surface excess. (From ref. 1.)

(FRES), is well-suited to tackle this problem for isotopic blends, because it permits the relative amounts of ^1H and ^2H to be measured as a function of depth.

In the technique, high-energy (about 3-MeV) helium ions are incident on the sample at a glancing angle. When these ions collide elastically with the ^1H and ^2H nuclei, the nuclei are ejected and subsequently detected; a foil screens the scattered He^{2+} ions from the detector. Kinematic theory predicts the energy with which nuclei at the surface are ejected (the energy separation between the two types of nuclei is about 0.5 MeV). For nuclei below the surface, energy is lost owing to frequent collisions with electrons, and so the peaks associated with the two types of ions are spread to lower energies, with a peak shape determined by the composition profile. The depth resolution of the technique (about 80 nm) is largely limited by the straggling in the stopper foil.

This technique was originally developed to look at ^1H and ^2H profiles in metal³. More recently Kramer's group have extended it to look at various diffusion problems in isotopic polymer blends. For the case of preferential segregation of one component at a free surface, the resolution is insufficient to reveal the detailed structure, but Kramer and colleagues can observe a marked enhancement in the average concentration of deuterated polystyrene relative to the bulk concentration (see figure). From the dependence of the surface excess on the bulk composition, the authors calculate, using an analysis developed by Schmidt and Binder⁴, that the difference between the surface energy of the protonated and deuterated forms of polystyrene is a mere $8 \times 10^{-5} \text{ J m}^{-2}$ — far too small to be measured directly.

But clearly, if such a weak effect in polystyrene (which arises from differences in polarizability between $\text{C}-^1\text{H}$ and $\text{C}-^2\text{H}$ bonds) leads to such a strong surface enhancement, then for polymer blends of practical importance, which typically contain components with very different chemical properties, the surface enhancement should be striking.

Because it is the blend's surface which

determines many of its properties — environmental resistance, wetting and friction, for instance — this finding is not simply of theoretical interest. In addition the surface properties may vary with the route of blend preparation — that is, whether this surface enrichment has time to become established.

The FRES technique is currently being refined by Kramer and co-workers, who are looking at ways the use of a stopper foil can be avoided to improve the resolution. They now claim² to have a detector with resolution better than 35 nm — capable of examining the composition profile in considerable detail. Other surface techniques with high spatial resolution are also becoming available for polymer systems. These include secondary-ion mass spectroscopy and X-ray photoelectron spectroscopy, but the one that looks most promising and is attracting most attention is neutron critical reflectometry. Many laboratories, including the Rutherford-Appleton, are now setting up

appropriate spectrometers (see ref. 5 for the first application to a polymer interface). Although this technique cannot provide such a direct measure of chemical composition as FRES, its high resolution (estimated to be about 5 Å) will permit checking of any hypothesized distribution of components. In this sense it can be regarded as complementary to FRES. A burst of activity in the field of polymer surfaces therefore seems likely, with implications for an improved understanding of surface phenomena at both the pure and applied levels. □

A.M. Donald is at the Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, UK.

1. Jones, R.A.L., Kramer, E.J., Rafailovich, M.H., Sokolov, J. & Schwarz, S.A. *Phys. Rev. Lett.* **62**, 280–283 (1989).
2. Sokolov, J., Rafailovich, M.H., Jones, R.A.L. & Kramer, E.J. *Appl. Phys. Lett.* **54**, 590–592 (1989).
3. Doyle, B.L. & Peercy, P.S. *Appl. Phys. Lett.* **34**, 811–813 (1979).
4. Schmidt, I. & Binder, K. *J. Phys., Paris*, **46**, 1631–1644 (1985).
5. Fernandez, M.L. *et al. Polymer* **29**, 1923–1928 (1988).

NATURAL SELECTION

Why do plants produce so many more ovules than seeds?

Deborah Charlesworth

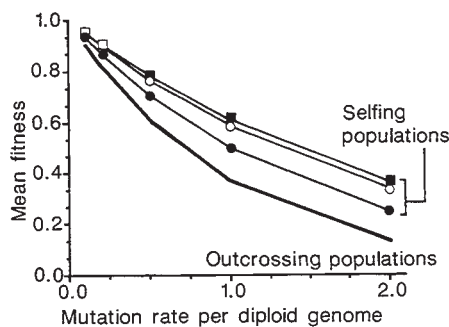
MANY flowering plants produce far fewer mature fruits than flowers, and fruits often contain fewer seeds than the number of ovules that were available for fertilization. A striking example of a plant in which there is a low value of the number of seeds produced compared with the number of available ovules is the shrub *Dedeckera eurekaensis* described by Wiens *et al.* on page 65 of this issue¹.

One obvious possibility, which seems to be at most only part of the story, is that flowers sometimes do not get enough pollen to fertilize all the ovules. The supply of pollen (or uncertainty of its receipt) can often limit the chance that a flower sets a full complement of seeds^{2,3}. In self-incompatible species, pollinators may deposit the plant's own pollen on its stigmas, perhaps clogging up the flowers' stigmas so that they cannot later be fertilized^{4,5}. Similarly, pollen from the wrong species might reduce flowers' chance of fertilization. In self-compatible species, ovules fertilized by self pollen may yield inviable seeds so that seed production in nature may be low.

In many cases, however, including that of *D. eurekaensis* described in this issue¹, pollen supply does not seem to be the main limiting factor for fruit set⁶ or seed set per fruit⁷. Because it is also frequently found that many fruits and seeds initiated are aborted even when plenty of compatible pollen is deposited on the stigmas^{8–10}, the role of the 'excess' flowers and ovules

cannot be to increase female fertility, and fruit abortion probably indicates limitations in the amount of resource available for the (presumably) costly process of fruit maturation¹¹. Under natural pollination conditions, outcrossing species generally have lower seed/ovule and fruit/flower ratios than do inbreeders, and so do long-lived woody species^{9,12,13}. The existence of these clear between-species differences suggests that these ratios are probably not just consequences of the local environmental circumstances in which the particular plants studied were growing. Interpreting such comparative data is difficult^{14,15}, in this case especially so because the characteristics of the different species are correlated: for example, annual species include a much lower fraction of outcrossers than do perennials¹⁵. Thus the low ratios of perennials might result from their more outbreeding population structure, compared with annuals, rather than from their living longer. The patterns seen can be greatly clarified by experimental data, which can help in the assessment of the contributions of the two most plausible types of explanation.

The first explanation is based on the idea from the theory of life-history evolution that perennial species must allocate resources between the competing functions of pollen and ovule production at the time of flowering (which probably determine later expenditures necessary to mature the fruits) and storage of nutrients



Mean fitness of a completely outcrossing population at its equilibrium under mutation to deleterious partially recessive alleles at many loci, compared with the corresponding mean fitness of populations reproducing only by self-fertilization, with different mutation rates per diploid genome. The outcrossing mean fitness does not depend on the type of selection or the values of the selection parameters, so only a single line is shown. For the selfing cases, results for various values of the selection coefficient of a mutant allele (s), and the dominance coefficient (h) are: black circles, $s=0.5$, $h=0.5$; open circles, $s=0.2$, $h=0.2$; squares, $s=0.01$, $h=0.2$.

for growth and survival to the next season; they are therefore expected to have a low rate of reproduction and thus produce few seeds per unit mass of plant in any one season¹⁶. Why then do they have many flowers and ovules? This can be accounted for in terms of the male fertility component of fitness through pollinator attraction^{17,18}. It is then surprising that plants do not frequently produce some flowers with purely male function and become andromonoecious. Total loss of female function of flowers may be disadvantageous because of uncertainty of the pollination process; unless each flower has a high chance of getting fertilized, or unless fruits are very costly of resources (for example very big), maximum-realized fertility cannot be achieved from a few flowers^{19,20}.

Another interpretation is that outcrossers suffer high genetic loads, causing many fertilization products to abort early in development²¹. (Genetic load is the relative lowering of the mean fitness of a population compared with the fitness of the best possible genotype.) Genetic loads are thought to result mainly from mutation to harmful partially or fully recessive alleles such as recessive lethals, and possibly also to alleles maintained by heterozygote advantage²². The second type of genetic load is least with outcrossing and increases somewhat with self fertilization (selfing), presumably because heterozygote advantage helps maintain variation in populations even though homozygotes produced by selfing are at a disadvantage. This type of gene action cannot therefore explain the low fecundity of outcrossing species.

With genetic load resulting from mutation, mean fitness increases as selfing increases because the equilibrium fre-

quences of deleterious alleles are reduced. The total genetic load in a wholly selfing population can be much lower than in a completely outcrossing population when the mutation rate is high, so that individuals in the outcrossing population are heterozygous for many mutant alleles depressing fitness (see figure). It seems unlikely that all this load will be expressed in the earliest stages of development so as to be detectable as reduced seed maturation, but this view might be wrong, as this is an important developmental stage. It is not known what proportion of plant genomes are expressed in the seed stage; certainly many mutations are known that affect seed (embryo and endosperm) characters^{23,24}.

If mutational load contributes to the loss of fertilization products in outcrossing plants, the abortion of individual seeds will be determined by the zygote genotypes; it should be random within fruits, and its frequency should be affected by the maternal and paternal parents, and by interaction between them²⁵. Furthermore, in outcrossing populations with mutational load including partially or fully recessive alleles, strong inbreeding depression would be expected; in *D. eurekaensis* Wiens *et al.* find that outcrossed flowers produce more mature seeds than selfed ones, but the difference is not statistically significant. This hypothesis could also account for the correlation between longevity and low female fecundity if long-lived species have higher mutation rates, as has recently been argued²⁶. □

Deborah Charlesworth is in the Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, USA.

- Wiens, D.C., Nickrent, D.L., Davern, C.I., Calvin, C.L. & Vivrette, N.J. *Nature* **338**, 65–67 (1989).
- Bierzychudek, P. *Am. Nat.* **117**, 838–840 (1981).
- Bertin, R.I. *Am. J. Bot.* **69**, 124–134 (1982).
- Bertin, R.I. *Am. J. Bot.* **75**, 1140–1147 (1988).
- Shore, J.S. & Barrett, S.C.H. *Can. J. Bot.* **62**, 1298–1303 (1984).
- Snow, A.A. *Oecologia (Berlin)* **55**, 231–237 (1982).
- McDade, L.A. & Davidar, P. *Oecologia (Berlin)* **64**, 61–67 (1984).
- Stephenson, A.G. *A. Rev. Ecol. Syst.* **12**, 253 (1981).
- Sutherland, S. & Delph, L.F. *Ecology* **65**, 1093 (1984).
- Mazer, S.J. *Evolution* **41**, 355–371 (1987).
- Stanton, M.L., Berczky, J.K. & Hasbrouck, H.D. *Oecologia (Berlin)* **74**, 68–76 (1987).
- Wiens, D. *Oecologia (Berlin)* **64**, 47–53 (1984).
- Sutherland, S. *Evolution* **40**, 117–128 (1986).
- Ridley, M. *The Explanation of Organic Diversity: the Comparative Method and Adaptations for Mating* (Oxford University Press, 1983).
- Felsenstein, J. *Am. Nat.* **125**, 1–15 (1985).
- Primack, R.B. *Am. Nat.* **114**, 51–62 (1979).
- Queller, D.C. *Nature* **305**, 706–707 (1983).
- Bell, G. *Proc. R. Soc. B* **224**, 223–265 (1985).
- Bertin, R.K. *Evol. Theor.* **6**, 25–32 (1982).
- Whalen, M.D. & Costich, D.E. in *Solanaceae: Biology and Systematics* (ed. D'Arcy, W.G.) 284–305 (Columbia University Press, New York, 1986).
- Wiens, D. *et al.* *Oecologia (Berlin)* **71**, 501–509 (1987).
- Charlesworth, D. & Charlesworth, B. (*A. Rev. Ecol. Syst.* **18**, 237–268 (1987).
- Mendel, G. *Verh. des Naturforsch. Ver. Brunn* **4**, 3–47 (1865).
- Meinke, D. *et al.* in *Plant Genetics* (ed. Freeling, M.) 129–146 (Liss, New York, 1985).
- Mazer, S.J., Snow, A.A. & Stanton, M.L. *Am. J. Bot.* **73**, 500–511 (1986).
- Klekowski, E.J. *Mutation, Developmental Selection, and Plant Evolution* (Columbia University Press, New York, 1988).

Surprising damp

DAEDALUS once devised an anti-condensation paint for kitchens and bathrooms. It was based on the 'hydrogel' polymers used for soft contact lenses; these readily absorb water, swelling somewhat as they do so. A hydrogel paint can thus absorb sudden peaks of condensation, and slowly breathe it out again as vapour in periods of low humidity. With the walls and ceiling of the whole room as an absorption surface, many litres of condensation can be taken up.

The problems began when Daedalus tried to colour his new product for the domestic market, by mixing it with a rather hastily formulated emulsion paint. In a cold, damp room the new mixed paint certainly took up water vapour very efficiently. But when the room was warmed up to reduce its relative humidity, condensation actually oozed from the painted surface and ran depressingly down the walls. It turned out that the liquid emulsion droplets dispersed in the paint had solidified in the cold test room. On warming up, they melted again, expanding strongly as melting vegetable oils do. The water-swollen hydrogel was thus suddenly and forcefully pressurized from within, and its loading of liquid water was squeezed out of it by reverse osmosis. Daedalus had invented a condensation generator.

DREADCO's marketing department soon rose to the challenge of this new product. Wet Paint® is being aimed at the car and vehicle market. The weekly wash and polish beloved by devout car-owners is mainly symbolic: corrosion starts at internal seams and surfaces no owner can reach. But a Wet-Painted car will never need to be cleaned at all. Every night, as the temperature falls and humidity rises, its surface will absorb water vapour; every day, as the temperature rises, it will weep the water out again as a liquid. Coming out of the paint surface itself, the exuded water will clean far more efficiently than any external spray. Dirt and road salt will simply be dislodged from beneath and floated away. Applied to trains, houses, bridges, and many other external structures, Wet Paint will keep grime, mould and graffiti at bay. The whole urban environment will be wonderfully smartened up.

Some technical problems remain to be solved, of course. During its pressure-exudation phase, Wet Paint tends to weep not only at its outer surface, but also at its inner one, thus pumping itself forcefully off the wall. Reliable bonding to substrate or undercoat is needed. On the other hand, DREADCO's cosmeticists hope to apply this principle in a new and powerful moisturizing cream. Unique among such preparations, it will take up moisture from the air and force it into the wearer's skin under many atmospheres of hydraulic pressure.

David Jones

The pulsar inside SN1987A

SIR—The discovery¹ of a pulsar inside the remnant of supernova 1987A is not in itself surprising. The initial neutrino burst^{2,3} suggested the formation of a neutron star, and the low-energy X-ray flux (6–16 keV) observed by Ginga⁴ has been interpreted as evidence for an active pulsar. Apart from fluctuations on time-scales of days, the soft X-ray flux has remained constant at about 6×10^{36} erg s⁻¹ since its first appearance in January 1988 (Y. Tanaka, personal communication). The harder emission, above 20 keV, has declined in the past few months⁵, in accordance with the radioactivity model⁶.

We have suggested that the soft X-ray flux could be synchrotron emission from a non-thermal nebula, fed by the remnant pulsar at a rate of about 10^{38} erg s⁻¹ in the form of relativistic particles and magnetic field^{7,8}. The emission could be detected if the supernova envelope had undergone early fragmentation, shortly after the explosion. The pulsar input would amount to a significant fraction of the present bolometric luminosity of the supernova, and deviations in the decay of the light curve should be expected: preliminary indications of an extra energy input may already be available⁸.

The finding of the pulsar lends strong support to these ideas. The properties of the pulsar, however, are rather surprising, especially its extremely short period and the small value of the surface magnetic field.

The strength of the magnetic field at the light cylinder radius, $\sim 2.4 \times 10^6$ cm, where field lines corotate with the pulsar would move tangentially at the speed of light, is bracketed by the requirements that the electromagnetic energy output be less than SN1987A's present bolometric luminosity, 2.1×10^{38} erg s⁻¹, and more than the pulsed optical luminosity, 6.4×10^{35} erg s⁻¹. Thus 2.3×10^6 G $< B < 4.2 \times 10^7$ G, corresponding to a surface field between 3.3×10^7 and 6×10^8 G if dipolar geometry is assumed.

The SN1987A pulsar thus belongs to the class of weakly magnetized, rapidly rotating objects for which a 'recycling'

scenario is accepted as canonical. Clearly, we have here first-hand evidence that at least some objects of the class are 'primordial'¹⁰. One may then wonder about the size of the millisecond pulsar population, because a new production mechanism has been found and since the radio surveys performed so far have not covered adequately the submillisecond range. Also, we note that, if pulsars such as this one are born frequently, their magnetic field cannot be amplified up to 10^{12} G later on by thermoelectric instabilities¹¹, because this would entail a huge, delayed release of rotational energy, for which we have no astrophysical evidence.

In the discovery observations a weak sinusoidal modulation of the pulsar frequency is noted. If confirmed and interpreted at face value as evidence for a binary companion, the modulation would imply a companion mass of $\sim 10^{-3}$ times the mass of the Sun. At any rate, whether the claimed modulation is real or only an upper limit, the interpretation that the soft X-rays from SN1987A result from accretion¹² appears unlikely.

Concerning the nature of the optical radiation mechanism, we note that the brightness temperature of the observed pulses, even including the entire speed-of-light surface, corresponds to an electron Lorentz factor of at least 1.6×10^4 , whereas the Lorentz frequency in the minimum magnetic field is already 6×10^{12} Hz. Then, in order to satisfy the thermodynamical constraint, one has to invoke either synchrotron radiation by protons, or by coherent processes, or, perhaps, electron synchrotron radiation well beyond the speed-of-light distance.

Another puzzle concerns the variations observed in the average pulsar optical luminosity. One could imagine a single physical cause for them and for the 8-h period of the frequency modulation, if the latter is confirmed. More likely, they could be attributed to the effect of the inhomogeneous nebular medium, in the same way as postulated for the fluctuations of the soft X-ray flux^{7,8}. Although homogeneous models of the supernova envelope become transparent to optical radiation a few months after the explosion¹³, the optical spectra of SN1987A still show clear signs of opacity effects, probably due to bound-bound and bound-free transitions of metals (E. Oliva, personal communication); if the envelope material is clumpy, and if the clumps have some non-radial motion, temporary occultations of the pulsar are a reasonable possibility.

Finally, we recall that a rotation frequency of 2 kHz is close to the hard limit of the virial theorem, and strains the predictions of certain stellar interior models on the onset of non-axisymmetric

instabilities¹⁴. The mere fact that such a rotation is observed after two years implies a mass quadrupole moment smaller than 5.3×10^{10} g cm². The above limit corresponds to a gravitational wave luminosity of 6.1×10^{14} erg s⁻¹, which still would be the main energy output channel by a very large margin. A measurement of the first and second derivatives of the period would establish the relative importance of gravitational radiation losses.

M. SALVATI

F. PACINI

R. BANDIERA

*Osservatorio Astrofisico di Arcetri,
Largo Enrico Fermi 5,
50125 Firenze, Italy*

Hair light guide

SIR—During studies on the staining of whole plucked hairs for use in biological dosimetry, we have observed that grey hairs illuminated at one end emit light at the other (J. Wells, *Stain Technol.* **63**, 189; 1988). An example of this phenomenon is illustrated in the figure. A 6-mm-long grey hair was mounted through a piece of card, ensuring that light could not



pass between the card and hair, and the card was placed on a microscope stage and illuminated from below. Of the 2 mm of hair protruding through the card, only the cut end emitted sufficient light to be seen down the microscope. Reduction in light intensity was noted with increased hair length but by far the most important factor was hair colour; little or no light was transmitted by brown hair.

Thus, grey hair acts as a natural fibre optic which can transmit light to its matrix, the follicular epithelium and to the dermis. Whether light transmission down hairs affects skin and hair needs investigation.

J. WELLS

*Radiobiology Laboratory,
Central Electricity Generating Board,
Berkeley Nuclear Laboratories,
Gloucestershire GL13 9PB, UK*

- Middleditch, J. et al. *IAU Circ.* No. 4735 (1989).
- Hirata, K. et al. *Phys. Rev. Lett.* **58**, 1490–1493 (1987).
- Bionta, R.M. et al. *Phys. Rev. Lett.* **58**, 1494–1498 (1987).
- Dotani, T. et al. *Nature* **330**, 230–231 (1987).
- Sunyaev, R. et al. *IAU Circ.* No. 4691 (1988).
- McCray, R., Shull, J.M. & Sutherland, D.P. *Astrophys. J.* **317**, L73–L77 (1987).
- Bandiera, R., Pacini, F. & Salvati, M. *Nature* **332**, 418–419 (1988).
- Bandiera, R., Pacini, F. & Salvati, M. *Astrophys. J.* (in the press).
- Burki, G. & Cramer, N. *IAU Circ.* No. 4729 (1989).
- Pacini, F. *Astr. Astrophys.* **126**, L11–L12 (1983).
- Blandford, R.D., Applegate, J.H. & Hornquist, L. *Mon. Not. R. astr. Soc.* **204**, 1025–1048 (1983).
- Fabian, A.C. & Rees, M.J. *Nature* **335**, 50–51 (1988).
- Bahcall, J.N., Rees, M.J. & Salpeter, E.E. *Astrophys. J.* **162**, 737–742 (1970).
- Lindblom, L. *Astrophys. J.* **303**, 146–153 (1986).

DNA helical repeats

SIR—The recent paper, "The helical repeat of double-stranded DNA varies as a function of catenation and supercoiling", by Wasserman *et al.*¹ gives the misleading impression that the authors have solved the fundamental problem of the partition function between twist and writhe in supercoiled DNA. In fact, the authors did not measure the DNA helical repeat, H (ref. 2), but a newly defined parameter, h (ref. 3).

Whereas H is uniquely determined by the three-dimensional structure of a given DNA molecule, h measures the relationship between the DNA and a chosen reference surface³, and is therefore related to DNA structure only indirectly. Thus, the differences in h calculated by Wasserman *et al.* correspond to changes in the relationship between the DNA and a reference surface rather than to changes in the internal structure of the DNA. The results of Wasserman *et al.* therefore have no relevance to the energetic constraints on DNA deformation. Furthermore, although the parameter h has a clear biological meaning when it relates DNA to a real surface, its significance becomes doubtful when the surface is only imaginary.

ANDRZEJ STASIAK

*Institute for Cell Biology,
Swiss Federal Institute of Technology,
CH-8093 Zurich, Switzerland*

CLAUDE WEBER

*Institute of Mathematics,
University of Geneva,
CH-1211 Geneva, Switzerland*

JACQUES DUBOCHET

*Centre of Electron Microscopy,
University of Lausanne,
CH-1005 Lausanne, Switzerland*

COZZARELLI ET AL. REPLY—For a DNA whose axis lies on a surface, real or virtual, there are two related measures of the turning of the double helix: twist (Tw), which measures a component of the rotation of one strand about the axis, and winding number (Φ)², which is the sum of the periodic exposures of either strand away from the surface. The helical repeat of a DNA with N base pairs has been defined in two different ways, namely (N/Φ) and (N/Tw) or h and H , respectively, in the nomenclature of Stasiak *et al.* Both are valuable measures of DNA structure and differ only when DNA is supercoiled.

Contrary to the assertion of Stasiak *et al.* however, h is often the property of biological significance. It is also much more easily measured. The parameter h defines the sequence periodicity of nucleosomal DNA (ref. 4) and the phasing of protein-binding sites, and is, in fact, most commonly measured by the accessibility of DNA to chemical or enzymatic probes. In our experiments, (N/Φ) not

(N/Tw) is determined by the catenane-induced supercoiling¹. Supercoiling derived from catenation, linking number deficit, and the wrapping of DNA around histones, in all cases, leads to a change in h (and also H); it is thus obviously incorrect for Stasiak *et al.* to assert that h has no relevance to DNA structure or deformation. Stasiak *et al.* make an additional error; Wang² did not distinguish H from h as he used relaxed DNA, for which the two are equal.

STEVEN A. WASSERMAN

NICHOLAS R. COZZARELLI

*Department of Molecular Biology,
University of California,
Berkeley, California 94720, USA*

JAMES H. WHITE

*Department of Mathematics,
University of California,
Los Angeles, California 90024, USA*

1. Wasserman, S.A., White, J.H. & Cozzarelli, N.R. *Nature* **334**, 448–450 (1988).
2. Wang, J.C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 200–203 (1979).
3. White, J.H., Cozzarelli, N.R. & Bauer, W.R. *Science* **241**, 323–327 (1988).
4. Satchwell, S.C., Drew, H.R. & Travers, A. *J. molec. Biol.* **191**, 659–675 (1986).

Cystic fibrosis

SIR—Kitzis *et al.*¹ confirm preliminary reports² of a segregation distortion of cystic fibrosis (CF) alleles with the sex of carriers. They use informative linked DNA markers to track CF alleles in families where there are carriers of the disease and count the numbers of carriers and non-carriers amongst unaffected siblings. The male/female ratio is as expected (1/1) and the carrier/non-carrier ratio follows mendelian expectation (2/1). But the male/female ratio for CF carriers is 1.21/1.0 (with a complementary 1.36/1 female/male ratio for non-carriers). We bring to your attention work reported by Gedschold *et al.*³, which we believe indirectly supports this segregation distortion phenomenon. These authors analysed questionnaires from a population-genetic study⁴, to show that for East German CF families, female carriers had more sibs (1.99 on average) than male carriers (1.66).

These two findings are remarkably consistent; the segregation distortion will generally lead to female carriers being found in larger sibships than male carriers. Using a male/female carrier ratio of 1.21/1.0, the likelihood ratio of families ascertained through one female carrier with 1.99 unspecified sibs producing n carrier sibs to families with one male carrier and 1.66 unspecified sibs producing n carrier sibs is almost unity (that is, ignoring constant terms, $1.21 \times 1.66/1.0 \times 1.99 = 1.0093$).

Hence, we conclude that the segregation distortion of CF alleles with sex explains the differences in sibship sizes for male and female carriers. Linkage dis-

equilibrium between the CF allele and alleles detected by DNA probes can be used to modify prior risks for individuals seeking carrier determination/exclusion⁵. Perhaps, now that the segregation distortion is confirmed, risk calculations should include the consultant's sex.

MARTIN FARRALL

PETER J. SCAMBLER

*Department of Biochemistry and
Molecular Genetics,
St Mary's Hospital Medical School,
Norfolk Place, London W2 1PG, UK*

1. Kitzis, A. *et al. Nature* **336**, 316 (1988).
2. Kitzis, A. *et al. Nature* **333**, 215 (1988).
3. Gedschold, J. *et al. Hum. Genet.* **80**, 399–400 (1988).
4. Gedschold, J. *Hum. Genet.* **75**, 277–280 (1987).
5. Farrall, M., Estvill, X. & Williamson, R. *Lancet* **ii**, 156–157 (1987).

Unhealthy genes

SIR—In a recent review article (*Nature* **336**, 435; 1988) Kondrashov asks whether future generations will gradually have less healthy genes. I found his article particularly interesting as I have recently started a debate in Norway concerning this problem.

The rate of mutation of human genes is no less than before. In industrialized societies, where the combination of medical care, lifestyles and reproduction strategies have decreased the selection for robust individuals, unwanted mutations will therefore be expected to accumulate.

The known genetic diseases represent only a small part of this problem, simply because they represent only a few specific mutations in a very limited number of genes. We have about 100,000 different genes (plus a large amount of DNA that influences gene activity) and each gene can mutate in many different ways. Some mutations are lethal, and therefore easily selected against, but most are probably either relatively neutral or slightly detrimental and do not cause any overt disease that can be traced to a specific mutation.

As a consequence we must, in the long run, expect an increased frequency of various health problems, implying that an increasing fraction of resources will need to be directed to health care.

I believe that an effort should be launched to evaluate the problem. The most important question is how fast the deterioration takes place. Depending on the outcome of this evaluation, it may be worth investigating possible means of dealing with the problem.

Kondrashov points out that the increase in unhealthy genes is a practically irreversible process. The irreversibility lies in the fact that we are dealing with humans and not animals, for which relevant selection programmes would be easy to implement.

BJØRN GRINDE

*National Institute of Public Health,
Geitmyrsveien 75,
0462 Oslo, Norway*

The rich and the irradiated

Daniel J. Kevles

Uranium Frenzy: Boom and Bust on the Colorado Plateau. By Raye C. Ringholz. W.W. Norton: 1989. Pp.310. \$18.95. To be published in Britain in June, £13.95.

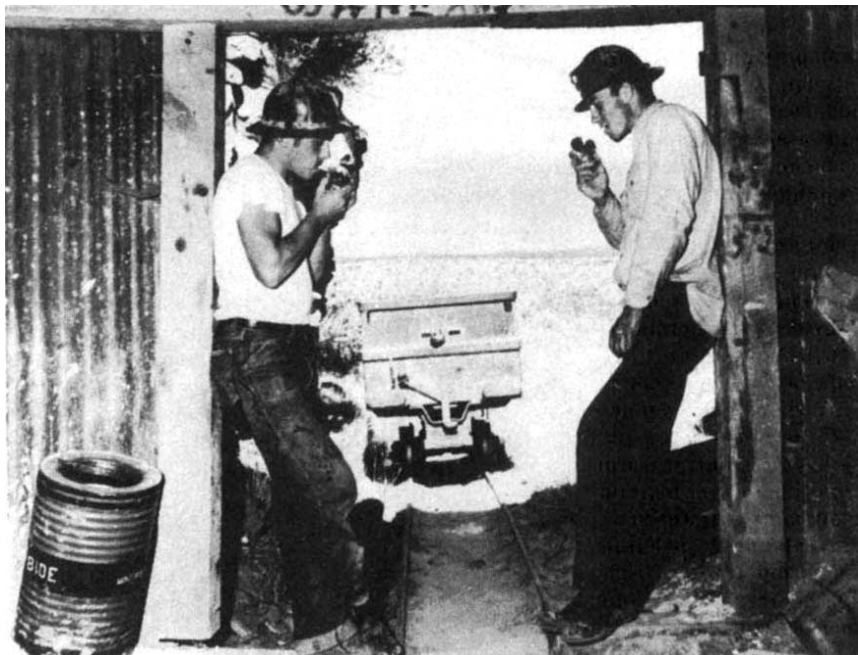
IN *Uranium Frenzy*, a tale of the modern American West, Raye Ringholz recounts the quicksilver dreams, leaden disappointments (and worse) brought about after the Second World War by the United States Atomic Energy Commission's desperate determination to encourage the establishment of a domestic uranium supply. The AEC, as the Commission was commonly known, offered prospectors a \$10,000 bonus for each discovery of a uranium lode and guaranteed to purchase ore production for ten years at a certain minimum price. The inducements helped to stimulate a boom in uranium prospecting, mining and production in the Colorado Plateau, the dry, unforgiving region where Colorado, Utah, Arizona and New Mexico conjoin.

The author is a longtime resident of the Plateau region and an experienced freelance author. She brings to her book a first-hand familiarity with the ambience of the boom years and information and anecdotes drawn from interviews — they form her principal sources — with many of the participants. Although attentive to the larger government policies that shaped the boom, she tells the tale by fastening on the stories of individual players, both winners (the prospectors and penny-stock jobbers who became millionaires overnight) and an array of losers. Some of the losers — the lucky ones — lost only their shirts. Others, found among the miners whose lungs were exposed to radon and its daughter products, came down, years later, with lung cancer.

Ringholz personifies the boom with Charlie Steen, who came to the Plateau, in 1949, as a young, rebellious and unemployed geologist to hunt for uranium. He searched for the ore in areas where conventional wisdom said it didn't exist and found a rich core sample of pitchblende when, in July 1952, he broke his last drilling bit at 197 feet. Charlie and his mother, Rose, who had grubstaked him, promptly formed the Utex Corporation to exploit the strike, joining with two partners, who put \$19,000 into the venture. In December 1953, Charlie and Rose bought their partners out, paying them \$3,272,000 for the Utex stock they had received for their \$19,000 investment slightly more than a year earlier.

By this time, the uranium boom was on in earnest. Moab, Utah, where Steen and his family had moved, had been a small, conservative, inbred town; now it pulsed with the traffic of Cadillac convertibles,

ore trucks and enormous Chevrolet station wagons. In 1952, the stock-salesman population of Salt Lake City had totalled 100 people working for 20 brokerages; by March 1955, it had reached 467 salesmen hawking their wares for 80 houses. The come-ons could be, to say the least, inventive. Squeeze the pitchblende, one



Double jeopardy — uranium miners take a cigarette break.

salesman suggested, deadpan, to a prospective customer who had mentioned that she suffered from rheumatism. After 15 minutes of squeezing, she said that she felt no more pain and bought \$50,000 worth of uranium stock. On May 27, 1954, seven million uranium shares changed hands in Salt Lake City, setting a daily record. A commentator noted in the *Salt Lake Tribune*, "It's a giant slot machine with uranium shares instead of bells" (p. 109).

Investors, high on paper profits one day, were brought to earth the next by fraud and bankruptcy. Utah officials resisted regulation by the United States Securities and Exchange Commission, preferring to keep oversight of uranium stocks in local hands, which were more sympathetic than federal authorities might be to venture capital fliers. Anyway, it was said, you couldn't legislate financial morals or protection. Eventually, many Utah brokerage houses were indicted for violations of federal security regulations and a number lost their broker's licences. In 1955, the boom in uranium stocks ended, but the uranium

mining continued. The next year, the first death from lung cancer occurred among the miners on the Plateau.

Ringholz makes clear that the federal government dealt rather less decisively with the hazards of radiation to uranium miners than it did with those of stock fraud to uranium investors. Not that federal officials were unaware of the danger to the miners: in 1950, Duncan Holaday, a member of the United States Public Health Service, had begun a study of radiation levels in the mines. By 1952, he had produced a report showing that the levels — a median in 48 mines of 3,100 picocuries of radon gas per litre of air — were far higher than those in European

mines where a large percentage of workers had eventually contracted lung cancer.

No copies of the report were made available to the miners — Holaday and his colleagues had gained access to the mines by tacitly agreeing not to disseminate their findings in a way that might alarm the men down in the shafts. High officials of the AEC declined to publicize it for fear that general knowledge of the risks might jeopardize the production of uranium. Before the report was produced, and also in the years after it, Holaday urged the adoption of various safety measures — notably much-improved ventilation — to reduce the radiation hazards. Some of the larger mines responded with meliorative action, but most of the operators on the Plateau did not. All the while, the AEC insisted that it had no authority to impose safety standards. When the issue came up in 1959, the Commission typically passed the buck to state health agencies, which did virtually nothing, and to the Public Health Service, which had no regulatory leverage and whose only recourse, as Ringholz puts it, was to "implore, exhort,

attempt to educate" (pp. 148–150).

By the early 1960s, the miners in Holaday's study cohort were dying of lung cancer at a higher rate than expected on the basis of the incidence of the disease in the general population. In 1963, the widow of one of them, Mrs Eola Garner, started to press a legal case for compensation. In 1967, J.V. Reistrup, a reporter for the *Washington Post*, brought the plight of the miners to national attention and Willard Wirtz, the Secretary of Labor, on the authority of an old statute finally acted to set federal safety standards in the uranium mines. By 1977, when Mrs Garner at last won a favourable settlement from the state of Colorado, conditions in the mines were much safer.

Uranium Frenzy is not a probing analysis of high policy, but it is a fine example of high and absorbing journalistic art. It graphically depicts the behaviour and attitudes in the uranium game on the Colorado Plateau, delineating in moving counterpoint the United States govern-

ment's material success against the ill fate of the men who dug the precious ore out of the rock. Ringholz's prose is detached and clear-eyed, understated yet unsparing in its assessments of the winners, the victims, the public health officials who tried to forestall the tragedy that befell the miners — and the bureaucratic postures that prevented them from succeeding. Her book makes utterly plausible the summary of the uranium craze that Stewart Udall, the one-time Secretary of the Interior, supplied her in a letter, in November 1987: "What is most poignant is that the losers were innocent victims — and the Atomic Energy big shots were 'patriots' who lied to protect what they conceived as the 'national interest'" (p. 262). The lies may have taken the form mainly of suppressions and omissions, but they were lies, fatal lies, nonetheless. □

Daniel J. Kevles is the J.O. and Juliette Koepfli Professor of the Humanities, California Institute of Technology, Pasadena, California 91125, USA.

In his own time

Mark Ridley

The Non-Darwinian Revolution: Reinterpreting a Historical Myth. By Peter J. Bowler. *The Johns Hopkins University Press: 1988. Pp. 238. \$30.25, £17.50.*

THE Reverend Thomas Malthus, who is well known for his belief that population growth must exceed the food supply, argued also that the resulting food shortage is beneficial for the human race. It raises people out of idleness and torpor. "The necessity of food for the support of life [he wrote] gives rise, probably, to a greater quantity of exertions than any other want, bodily or mental." It is these exertions that have been the motor of human progress — "Had population and food increased in the same ratio, it is probable that man might never have risen from the savage state". Malthus's ideas would develop into the familiar victorian idea of 'progress through struggle': we desire to better our condition, we strive to do so and (provided the government follows a *laissez faire* policy) progress is the result.

Historians, and biologists, have often described Darwin's theory as an expression of the broader victorian political economy. Strictly, this has no bearing on whether Darwin was right; but Darwin's critics have always been fond of the analogy, both for the ideological tarnish and the chance of compromising Darwin's originality. Peter Bowler's latest book is most interesting for the thorough, critical look it takes at the historical question.

Bowler has to deal with three argu-

ments. First, that Darwin's theory apparently resembles those of Malthus, and of Herbert Spencer and other victorian writers; second, that natural selection was simultaneously discovered, at least by Alfred Russel Wallace and perhaps by Patrick Matthew and Edward Blyth, which suggests that the idea was indeed in the air; and finally, that Darwin's theory was rapidly accepted and developed into the influential politics of social darwinism.

Bowler has objections to all three arguments. For a start, the way that the struggle for existence led to progress was crucially different in the theories of Darwin and of the political economists. For Malthus, and Spencer, the struggle worked as a stimulus to individual action, whereas in Darwin's theory selection takes place on inherited differences between individuals within a population. As Bowler neatly puts it, in Spencer's lamarckist theory, evolution takes place in individuals, not populations; it lacks what Ernst Mayr calls 'population thinking'.

Wallace's theory may also have differed from Darwin's. In modern population genetical terms, Wallace appears to have been a group selectionist who thought of all selection as 'hard'. He was concerned with selection between "varieties", not individuals, and thought that selection only took place when a variety did not possess an adaptation needed to survive the external environment. Wallace's theory, like Spencer's, lacked Darwin's ideas of intraspecific competition and individual selection.

Finally, Darwin's theory was hardly accepted at all in the late nineteenth, or the early twentieth, centuries. What actually happened was that Darwin persuaded

biologists only about evolutionary change. They then called themselves darwinians, but were really (what Bowler calls) "pseudo-Darwinians"; they believed not in the darwinian sort of branching, contingent, unplanned evolution, but in evolution through a series of developmental stages from protozoa to human beings. The developmental model of evolution resulted in the recapitulatory, haeckelian kind of research in embryology, and in the theory of orthogenesis in palaeontology: both of these theories were still teleological, even if the static teleology of pre-darwinian natural theology had become the temporal teleology of orthogenesis. Bowler also argues that social darwinists were inspired by Spencer's lamarckist and progressive model of evolution, rather than by Darwin. In another chapter he describes a similar tendency in anthropology, as human evolution was considered to proceed through a series of racial stages, from savages to civilized Westerners.

Darwin thus appears as a historically isolated figure. There was no 'darwinian revolution'. The main stream of victorian thought started with the natural theologians and idealists, such as Richard Owen, who believed in a more or less atemporal plan of nature. It ran through Chambers, and Spencer, who tried to temporalize the plan, but proved unpersuasive. Darwin, however, was successful. He unintentionally led late-nineteenth-century biologists to accept a temporal, though still teleological, plan of nature, in which evolution proceeds through a series of stages, driven by some process of directed variation. He convinced no one either about his branching and non-progressive theory of evolution, or his materialist mechanism for it.

The 'Darwin industry' has often been criticized for concentrating on Darwin and ignoring his historical context. Bowler himself has some sharp things — not well directed, in my opinion — to say about that industry. *The Non-Darwinian Revolution* is a professional and worthwhile book, which merits reading; it offers a new (at least for the non-specialist reader) interpretation of events about which most people have opinions. I only really disagree with the way Bowler treats his fellow historians, whom he from time to time ignores or caricatures, but this is only a minor defect. Indeed, the industry itself may enjoy the last laugh. If we look past Bowler's various criticisms, and concentrate on his conclusions, he appears to have vindicated those whom he purports to oppose. Bowler has given Darwin his real victorian context: and ends up justifying those who study Darwin in isolation. □

Mark Ridley is in the Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK.

Into the fishy business

J. H. S. Blaxter

The Provident Sea. By D. H. Cushing. Cambridge University Press: 1988. Pp. 329. £37.50, \$65.

THE history of fisheries demonstrates some of mankind's better individual qualities — the ability to innovate and organize, and ruggedness under harsh conditions — as well as some less-desirable national foibles such as aggression and greed. These human attributes are well brought out in Cushing's book, which shows that fishermen and fishery biologists have some of the better qualities in common!

The history of fishing is dealt with on a wide geographical basis and some fisheries (mainly of the Northern Hemisphere) are described case by case — for example the North Sea herring fishery, the Newfoundland cod and harp seal fisheries, the Pribilof fur seal fishery and the great fisheries for whales. The origins of fisheries research and its increasing sophistication between 1945 and 1965, and from 1965 to the present time, are examined in relation to similar phases in the development of the regulating bodies or institutions concerned.

In prehistoric times, fish were caught by hook and harpoon; nets were already in use in the Neolithic. But a well-documented expansion took place in high-seas fisheries in the Middle Ages when fish caught in the Baltic and North Sea became a major source of employment and wealth, for example for the Hanseatic towns. Soon a high degree of international organization and integration evolved, not only in the fishing itself, but in landing, on-shore processing, transport and marketing.

Certain advances can then be identified, each causing a new leap forward. The first was the invention of on-board curing of herring by the Dutch in the fourteenth century. There followed the introduction of steam tugs for towing sailing vessels in and out of port in the middle 1800s, and then steam engines for propulsion and handling gear on the fishing vessels themselves in the late 1800s. The post-1945 development of stern trawlers, freezing at sea, use of midwater trawls and purse seines, and the deployment of fishing fleets with mother ships and catchers increased the world catch enormously, leading to its probable present asymptote near 80 million tonnes a year.

A large part of the book consists of chapters on the origin of fisheries research in the last century and its progress after the Second World War. The concomitant setting-up of international regulatory bodies closely parallels the development

of research. Clearly, the initiative for research and regulation stemmed from the increasing power and efficacy of the fleets, which began to cause severe depletion of fish stocks, a problem of little importance in the pre-steam days.

As the author points out, at one time nations went to war over their fisheries; now they try to resolve their problems by appraising the scientific evidence. The story of the interplay between fishermen and fishing, and fishery scientists and regulation, is a fascinating mix of ignorance, greed, optimism, political expediency and idealism. Apart from some Scandinavian countries such as Iceland and Norway, where scientists may be held in some esteem, fishery biologists have been treated with suspicion by the fishermen, who have resented the loss of short-term profits and ignored the goal of long-term consistency of catch. The ever-increasing effect of overfishing has recently led to a greater appreciation by fishermen of the need for regulation. Earlier methods of regulation, such as minimum mesh sizes, have of late been supplemented by the imposition of quotas for vessels and total allowable catches for fisheries. Usually the medicine prescribed can easily

be agreed by the scientists on an international basis — even if the politicians find it difficult to allocate the dose to their respective countries.

Cushing does not commend his book to any particular readership. Parts of it are probably too technical for the historian or economist, or too specialized for the undergraduate. Rather, it seems ideal for MSc students on fisheries management courses, or as background reading for marine and fishery biologists as well as civil servants and politicians — throughout the world, fisheries are often now as much a political as a commercial or scientific matter. The author does not refer to the increasing importance of freshwater fisheries and aquaculture (together yielding about 11 million tonnes in 1986), nor does he attempt to forecast the future. Although such additions would have enhanced the value of the book, it already contains an unusual and useful mixture of history, economics and science. It will no doubt find its way into many libraries — not least, I hope, those of certain government departments. □

J. H. S. Blaxter is Deputy Chief Scientific Officer at the Dunstaffnage Marine Laboratory, PO Box 3, Oban, Argyll PA34 4AD, UK.

The Great Dane

H.B.G. Casimir

Harmony and Unity: The Life of Niels Bohr. By Niels Blaedel. Science Tech Publishers, 701 Ridge Street, Madison, Wisconsin: 1988. Pp. 323. \$35. Distributed outside North America by Springer-Verlag, DM98, £32.

NIELS Bohr was one of the greatest physicists of our century. His papers on the spectrum of hydrogen, in which he postulated the existence of stationary states and of quantum jumps, inaugurated the era of the quantum theory of line spectra, led to an understanding of the periodic system of the elements — another notable contribution of Bohr's — and finally to the formulation of quantum mechanics. Bohr did not contribute to the mathematical formalism of the new mechanics, but his doctrine of 'complementarity' played an essential role in clarifying its interpretation. He then turned from atomic to nuclear physics, introducing the notion of highly excited intermediate states in nuclear reactions and explaining many features of uranium fission. Through his writings, and even more through personal discussions, he had an enormous influence on younger physicists.

During the Second World War Bohr escaped from his native Denmark, reaching the United States by way of Sweden and England. He was keenly aware of the

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Niels Bohr — an enormous influence on younger physicists.

decisive role nuclear weapons were going to play in international politics, and foresaw with great clarity the threat of a cold war, but his attempts to convince Roosevelt and Churchill that they should come to terms with the Soviet Union at an early date were unsuccessful.

He was certainly a true citizen of the world. But, as one of his biographers, quoted by Blaedel, remarks, he was so deeply rooted in Danish culture and the Danish way of thinking and feeling that, if he had grown up in some other country, he would not have been Niels Bohr. He was also the head of a wonderful family. ►

There is no lack of published material about Bohr. The edition of his collected papers also contains many biographical details, letters and so on, and there have been numerous obituaries and at least two commemorative volumes. In this book, a translation and slight revision of the Danish edition of 1985, Blaedel has addressed himself to the task of writing a full-length biography that covers all facets of his subject and that emphasizes that they form part of one harmonious unity. I think that on the whole he has succeeded remarkably well. He gives an accurate picture of the man theorists of my generation both admired and loved. And not only of the physicist: Bohr's relations with his family and in particular with his wife, an admirable woman, are drawn with sympathy and understanding.

Blaedel's sketch of the atmosphere at Bohr's institute in Copenhagen, where work was combined with — often somewhat childish — playfulness, especially in the late 1920s and early 1930s, is true to life; it will raise nostalgic memories among those who, like myself, experienced it. I am not so sure about the physics, however. The essential subjects are included and I have not spotted obvious errors, but I wonder whether Blaedel's account of the discussions between Bohr and Einstein — to give one example — will be comprehensible to non-physicists. Here, though, is a book to be read as a biography and not as an introduction to modern physics.

The translation from the Danish is in general satisfactory, although I noticed one or two slips. A few errors in the original have been corrected, and here and there short explanations have been added to help readers who are unfamiliar with Danish geography and Danish history and literature. The book is richly illustrated; although the quality of reproduction is not as high as that of the rather sumptuous original (which also contains a few colour plates), it is perfectly adequate.

Some readers may regret that the book is more a eulogy than a critical appraisal — that Blaedel does not mention any scientific shortcomings or any less desirable human traits. So be it. Personally I think he has produced a fitting tribute to a great scientist and a noble man. □

H.B.G. Casimir, De Zegge 7, 5591 TT Heeze, The Netherlands, studied theoretical physics at Leiden, Copenhagen and Zürich, and from 1946 to 1972 was Research Director of the Philips Company at Eindhoven.

New in Britain

Just published by Unwin Hyman, price £11.95, is "What Do You Care What Other People Think" by Richard Feynman as told to Ralph Leighton. The book, which was published last autumn in the United States by W.W. Norton and reviewed in *Nature* on 3 November, tells largely of Feynman's productively unconventional work on the inquiry into the failure of the *Challenger* space shuttle.

Wholesome ideals

P. W. Atkins

Science, Order and Creativity. By David Bohm and F. David Peat. *Routledge: 1988. Pp.280. Pbk £5.95.*

THE world will be much nicer once 'creativity' has been liberated from its present tomb, the walls of which are built from the attitudes of conventional science. This is the message that David Bohm and F. David Peat urge on us in this somewhat bizarre and certainly uneven book that has grown out of conversations that the two have had together over the years. I began to read with high hopes. But although some of the early passages show interesting insights, in general my dislike grew with the page numbers.

The book is pervaded on the one hand by a schoolboy naïveté ("How does science intend to end the danger of mutual annihilation that exists in the world?") and on the other by a pessimism about the potential of and the attitudes currently struck by conventional science. Before considering purchasing it, those who are sensitive to these things will want to know that both Sir Fred Hoyle and Rupert Sheldrake are mentioned without disapproval.

To be fair, I must allow the authors to speak for themselves, for I fear I may be one of those at whom is directed the accusation of being a prisoner of conventional science, and therefore may be over-tender to its message. The argument begins with an attempt to disarm criticism of what is to come by considering the role of communication in creative perception, and by identifying how communication may break down at the dawn of revolutions in science and at what the authors so fervently long for, the sunrise of a new paradigm. Quantum theory is in this context a paradigm of paradigms, and something of the wars of interpretation is presented here. Rather inappropriately, though, Bohm takes the opportunity to present a moderately lengthy account of his 'causal interpretation', for which he needs the burden of a new potential (the quantum potential, essentially $\nabla^2\psi/\psi^2$). The advantage of his formalism appears to be that it does not require a break with classically established concepts; here, however, the authors have tripped over their own feet, for a more relaxed view of quantum theory would lead us to believe that we must unburden ourselves of the expectation of the farmyard and not cling to classical concepts.

Built on these foundations, such as they are, is an account of what is meant by order, its classification and (an essential feature of what is to come) its unfolding.

The authors deal properly with the change in our perception of the concept that has been stimulated by developments in fractals and determinate chaos. To do so, they distinguish between 'constitutive order' (the actual order expressed in the physical constitution of the entity, as in the material hexagons of a beehive) and 'descriptive order' (the order that stems from an abstraction, as in the motion of an object through a field of coordinates). They argue that actual order (whatever that is) lies between these two extremes, and thus alight on what is for them the comforting view that order is a cycle of activity that involves both. It is in this chapter that the generally 'Eastern' flavour of the book begins to be explicit for the first time.

The authors are desperate to eliminate the fragmentation of conventional science and thus to impress upon us the essential 'wholeness' that is crucial to comprehension and human progress. Hence order, of various degrees, kinds and hierarchies, is essential to their argument. Thus they go on to distinguish 'generative order' from 'implicate order'. The former is an "inward order out of which the manifest form of things can emerge creatively". The rule for generating fractals would be an example. The latter is a particular kind of generative order that, I must confess, leaves me confused, for to comprehend it we must somehow understand that "the basic movement of enfoldment and unfoldment is thus a dual one in which there is ultimately no separation between enfoldment and the unfoldment". There is, of course, a superimplicate order too.

In the tacit infrastructure of consciousness, the authors claim, there are rigid ideas that restrict creativity. (I suspect, on the contrary, that they constitute a ladder to comprehension.) These ideas operate close to the source of the generative order and, as well as restricting creativity, imply "the presence of an energy that is directed toward general destructiveness". Whatever that means, I feel that it is the manifestation of the authors' pessimism, and more truly the spur of their writings than the overt joy of enlarged comprehension that they claim to be the book's impetus.

Bohm and Peat urge on us the "key importance of liberating creativity" if human life is to have a worthwhile kind of survival. And how do we escape from our prison? We need to de-fragment science, attain wholeness, comprehend levels of order, give up the hope of true and total comprehension, and, perhaps most important of all, "clear up" the sociocultural dimension. It really is as easy as that. □

P.W. Atkins is in the Physical Chemistry Laboratory, University of Oxford, and is a Fellow of Lincoln College, Oxford OX1 3DR, UK.

Geochemical implications of the formation of the Moon by a single giant impact

Horton E. Newsom & Stuart Ross Taylor

The origin of the Moon by a single massive impact of a body slightly larger than Mars with the Earth can explain the angular momentum, orbital characteristics and unique nature of the Earth–Moon system. The density and chemical differences between the Earth and the Moon are accounted for by deriving the Moon from the mantle of the impactor. A cosmochemically plausible impactor can be formed in the region of the inner Solar System, lending support to the impact hypothesis.

THE Earth–Moon system has several unique features. None of the other terrestrial planets possesses a comparable moon—the tiny martian moons, Phobos and Deimos, are probably captured asteroids. The angular momentum of the Earth–Moon pair is anomalously high compared with that of the other inner planets, and the inclination of the lunar orbit is strange. The Moon has a high mass (1/81.3) relative to the Earth when compared with the satellites of the giant planets, but its bulk density (3.34 g cm^{-3}) is much lower than that of the Earth or the other inner planets, probably because of a low metallic iron content.

The chemical composition of the Moon, revealed by the samples from the Apollo and Luna missions, is also unusual. The Moon is bone dry, is strongly depleted in volatile elements such as potassium, lead and bismuth, and probably enriched in refractory elements such as calcium, aluminium, titanium and uranium (see box). In bulk, it contains ~50% more FeO than the upper mantle of the Earth.

In view of this diverse set of properties, it is perhaps not surprising that theories for the origin of the Moon have generally been unsatisfactory. Only recently has a model been proposed that is widely considered to be acceptable: the single impact hypothesis. This theory proposes that during the final stages of the accretion of the terrestrial planets, a body somewhat larger than Mars collided with the Earth and spun out a disk of material from which the Moon formed. Here we will review the evidence for this theory and show how it can account for the unusual chemical composition of the Moon.

Older hypotheses

Including the single impact theory, there are five main classes of hypothesis for the origin of the Moon, although elements of some appear in others. The four older hypotheses fail to explain the unique nature of the Moon because they do not account for the lunar orbit or the high angular momentum of the Earth–Moon system, and because they entail processes that might have been common in the early Solar System, implying that similar satellites should accompany all the inner planets.

Capture of an already formed Moon from an independent orbit is improbable on dynamical grounds¹ and does not explain

the compositional peculiarities, as the Moon would be expected to be an example of a common early Solar System object².

Formation of the Earth–Moon system as a double planet immediately stumbles on the difficulties of density and composition. To overcome the density problem, co-accretion was proposed³; in this model the Moon formed from a ring of low-density silicate debris shed from the mantles of incoming differentiated planetesimals, whose iron cores survived to accrete with the Earth. Although attractive, this scenario is improbable⁴ because the breakup of planetesimals is unlikely to occur and it is difficult to achieve the required angular momentum⁴. In a similar model for co-accretion and evolution of a circum-terrestrial disk^{5,6}, it has been shown that the angular momentum is attained only in very special conditions.

Fission hypotheses, which derive the Moon from the terrestrial mantle⁷, have been popular because they explain a low-density, metal-poor Moon, but they had to be abandoned after lunar samples showed significant differences in chemical composition between the Moon and the Earth's mantle⁸. A modification of this hypothesis^{9,10} suggests that mantle material was thrown into orbit by many small impacts. Apart from the inherent chemical difficulties, it is exceedingly difficult to obtain the required angular momentum this way¹¹. A further elaboration of this proposal, involving capture of objects in heliocentric orbits by an extended Earth atmosphere⁹, is not supported by existing evidence^{8,12,13}, and this model also fails to account for the unique status of the Moon.

Single impact hypothesis

Formation of the Moon as the result of a single collision of a large body with the early Earth resolves many of these problems. The theory was first proposed to account for the anomalous angular momentum¹⁴, but also provides an explanation for the other properties and has now become widely accepted¹⁵.

In assessing the single impact theory, one must first ask whether a suitably large body existed in the early Solar System (before 4.4 Gyr). One scheme for the origin of the terrestrial planets¹², in which they are accreted from a suite of rocky planetesimals over a period of 10^7 – 10^8 yr, leads to a hierarchy of large bodies. In the final stages of accretion, perhaps 100 objects of lunar mass, 10 of the mass of Mercury and a few Mars-sized bodies populate the inner Solar System¹² before being swept up into the four inner planets.

There is considerable evidence for this scheme. Craters and ringed basins over 1,000 km in diameter on the Moon, Mercury, Mars and the satellites of the outer planets¹⁶ attest to an early (>3.8 Gyr) intense bombardment by a large range of objects. The axes of nearly all the planets are significantly tilted relative to the plane of the ecliptic; the most dramatic example is Uranus, which is lying on its side, probably as a result of a collision with an Earth-sized object. The slow backward rotation of Venus, unique in the Solar System, is most rationally attributed to a

Classification of the chemical elements

To describe the behaviour of the elements in the solar nebula, cosmochemists divide them according to their volatility (examples in parentheses): 'gaseous' (hydrogen, carbon, nitrogen, oxygen and the noble gases), 'very volatile' (bismuth, thallium), 'volatile' (rubidium, caesium), 'moderately volatile' (potassium, manganese), 'moderately refractory' (vanadium, europium), 'refractory' (calcium, aluminium, uranium, lanthanum) and 'super-refractory' (zirconium, scandium). The terms lithophile, chalcophile and siderophile describe those elements that preferentially enter silicate, sulphide or metal phases, respectively. This classification is based primarily on the distribution of elements in these phases in meteorites.

late collision with a massive, perhaps Mars-sized object; with a different mass, angle and velocity, the impact might have provided Venus with its own moon.

Further evidence comes from the varied compositions of the planets; accretion from a dusty nebula, rather than from planetesimals, might be expected to produce rather uniform planetary compositions. Meteorites, too, come from many distinct parent bodies and are commonly mixtures of several components, or fragments of larger bodies. The asteroid belt, arrested at an early stage of planetary development, probably owing to the gravitational influence of the gaseous giant Jupiter, preserves a wide size range of objects (the largest, 1 Ceres, is 1,020 km in diameter). Finally, Mercury's anomalously high density and large iron core are most simply explained by removal of most of its silicate mantle by collision with a large body¹⁷; alternative schemes involving high-temperature evaporation of silicates remove unrealistic amounts of material¹⁸ and do not account for the presence of sodium ions, probably sputtered from the surface, in Mercury's tenuous atmosphere¹⁹.

Impact dynamics

Studies of the single impact hypothesis^{20–26} have led to the following sequence of events (see Fig. 1). When the Earth had attained nearly its present size, it suffered a grazing impact, at about 5 km s^{-1} , with an object of ~ 0.14 Earth masses, somewhat larger than Mars^{20,21}. Both this body and the Earth are assumed to have differentiated into a metallic core and silicate mantle. The collision disrupts the impactor, much of which goes into orbit around the Earth, accelerated by the gravitational torques arising from the asymmetrical shape of the Earth following the impact^{20,21}, and by expanding gases from the vaporized part of the impactor^{14,23}. While the impactor's mantle is being accelerated away from the Earth, its metallic core separates from the mantle, decelerates, and accretes to the Earth in about four hours²².

The material that achieves orbit is initially present as a disk partly inside and partly outside the Roche limit (~ 2.9 Earth radii)^{22,25,26}. Some of the material inside the Roche limit ends up outside through transfer of angular momentum. The material may immediately coalesce to form a totally molten Moon, or break up into several moonlets which then accrete to form a partly molten Moon²⁶. Geochemical studies indicate that at least half the Moon was molten shortly after accretion, with the feldspathic highland crust crystallizing from this 'magma ocean'²⁷; in general, a partly rather than fully molten Moon seems most consistent with geochemical and geophysical constraints²⁶.

Lunar composition

Only a small amount of material from the Earth's mantle (probably less than 16%^{20,22}) eventually ends up in the Moon. This is supported⁸ by the low FeO content of the terrestrial mantle (8%) as compared with the Moon (13%), which would otherwise require that the FeO content of the impactor mantle be unreasonably large ($>13\%$)²⁸. A Moon formed from 85% impactor mantle and 15% terrestrial mantle requires an impactor mantle with 14% FeO.

Additional evidence that there is relatively little terrestrial mantle material in the Moon comes from the isotopic composition of potassium²⁹. Because the Moon is depleted in volatile elements relative to the Earth's mantle, the mass fractionation due to volatile loss should have left the Moon enriched in the heavier potassium isotopes. The fact that the isotopic composition of potassium in lunar feldspars is close to that of CI carbonaceous chondrites, considered to be typical of the solar nebula²⁹, limits the contribution from the Earth's mantle to $<20\%$.

Relative to chondritic abundances (see Fig. 2 legend for a discussion of the use of carbonaceous chondrites as 'benchmarks'), the Moon and the Earth's mantle show similar depletion

patterns of vanadium, chromium and manganese^{9,30}. These contrast with the more chondrite-like abundances inferred for the parent bodies of the eucrite and shergottite meteorites, and are considered by some^{9,10} to support an origin of the Moon from the Earth's mantle. But the depletion of Mn relative to V and Cr in the Earth and the Moon is also consistent with volatility having been the primary control on the abundance pattern^{31,32}. New experiments on partitioning of these elements between sulphur-rich metallic liquid and silicate melt^{31,32}, confirming earlier work on partitioning between pure iron metal and silicate melt^{33,34}, as well as data for Cr and Mn in natural metal-silicate rocks from Disko Island³⁵, show that V and Cr are more highly siderophile (literally, 'iron-loving'; see box) than Mn. Thus, had these elements been depleted by core formation in the Earth, we would see a greater depletion of V and Cr compared to Mn, contrary to observation. Their depletion in the Earth's mantle, therefore, cannot be ascribed to unique terrestrial processes involving formation of the Earth's core^{9,36}, and similar depletions of V, Cr and Mn in the Earth and Moon do not require that the Moon formed primarily from terrestrial mantle material.

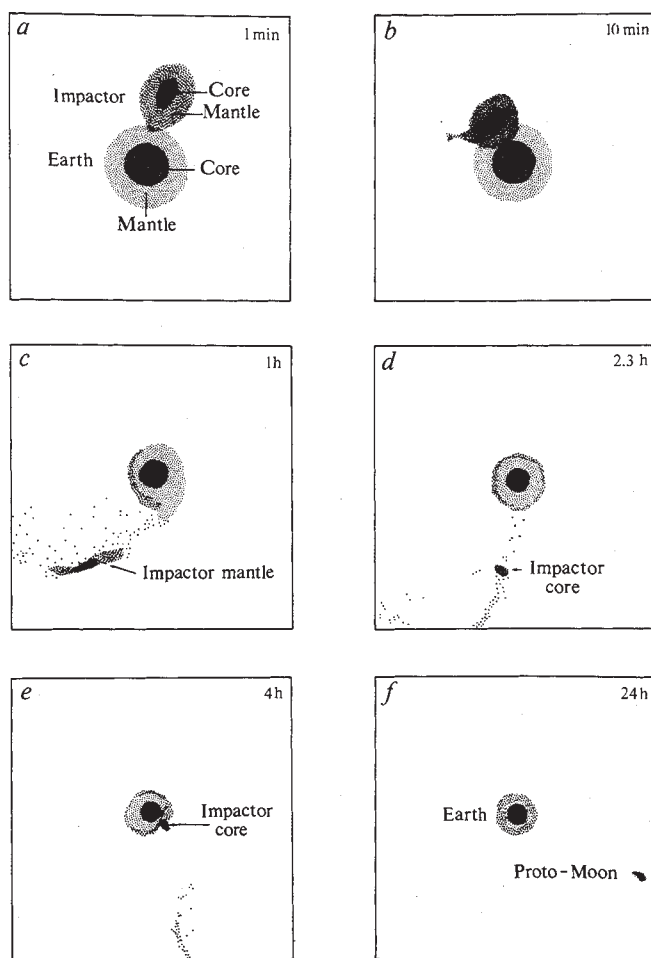


FIG. 1 Computer simulation^{20–22} of the formation of the Moon by a giant impact. This reconstruction shows the events following the oblique collision with the Earth of an object of 0.14 Earth masses (slightly larger than Mars) at a velocity of 5 km s^{-1} . Both the Earth and the impactor have already differentiated into a metallic core and silicate mantle. The time elapsed since impact is given in each box. Following the collision (a, b), the impactor is spread out in space (c), but the debris clumps together through gravitational attraction. The iron core of the impactor separates from the silicate mantle (d) and accretes to the Earth (e) about four hours after the initial encounter. Nearly 24 hours later (f), a silicate lump of about lunar mass is in orbit around the Earth. This material is principally derived from the silicate mantle of the impactor. Courtesy of A. G. W. Cameron and W. Benz.

Constraints on impactor composition

What can be deduced about the composition of the impacting body? We will first consider the constraints on the lithophile and volatile element composition of the impactor, and then use the siderophile element depletion patterns in the Earth and Moon to derive the siderophile abundances in the impactor's mantle and the size and composition of the impactor's core. We assume²⁰⁻²² that the impactor was 0.14 Earth masses, and was differentiated into a metallic core and silicate mantle. We also assume that it formed in the same part of the solar nebula as the Earth (between Venus and Mars), to account for the low relative impact velocity (5 km s^{-1} ; ref. 23) and the similarity in oxygen isotopes between Earth and Moon. Although there was probably much mixing of planetesimals during their formation, the terrestrial planets retain some memory of having accreted

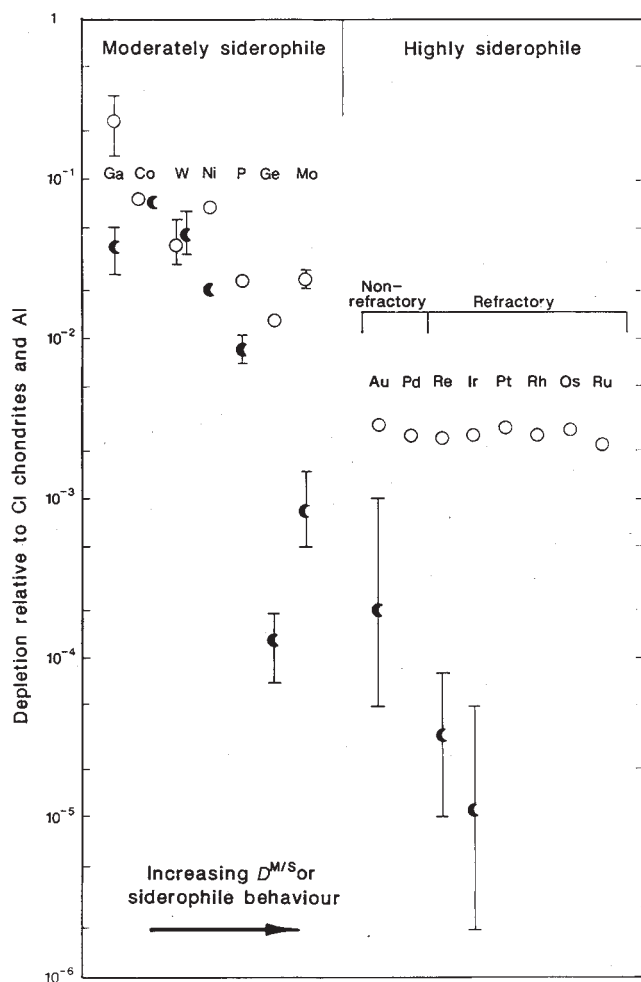


FIG. 2 Depletion of siderophile elements in the Earth (○) and Moon (●)^{42,51}, plotted in order of metal/silicate partition coefficient, $D^{M/S}$. The dividing line between moderately and highly siderophile behaviour is taken to be $D^{M/S} = 10,000$, meaning that the element is partitioned into the metallic phase with a concentration ratio (metal:silicate) of 10,000:1. The depletions are normalized to abundances in CI chondrites relative to aluminium, because the non-siderophile refractory elements are enriched in the Earth and Moon. The greater depletion of gallium and germanium in the Moon relative to elements with similar siderophile behaviour is probably due to their volatility. Elemental abundances in Solar System objects are usually expressed relative to those in the primitive, volatile-rich meteorites known as CI chondrites. For the non-gaseous elements, these meteorites give the best estimate of the composition of the Sun, and thus of the primitive solar nebula from which the planets formed⁵⁶. Primitive material, unchanged since the formation of the solar nebula, would have flat or 'chondritic' patterns; conversely, any deviation from chondritic abundances (or relative abundances) provides information about an object's chemical history.

from relatively narrow ($<1 \text{ AU}$) feeding zones, a notion supported by the zoned compositional structure of the asteroid belt³⁷. **Lithophile and volatile elements.** The constraints imposed by these elements on lunar origins and the single impact hypothesis are only briefly summarized here; for more details, see ref. 8. The impact was sufficiently energetic to vaporize much of the material that went to make up the Moon, explaining such unique features as the bone-dry nature of the Moon and the extreme depletion of very volatile elements. Bismuth and thallium, for example, are depleted by factors of ~ 200 relative to cosmic abundances. Elements that are volatile at temperatures above $\sim 1,100 \text{ K}$ (such as europium and ytterbium) do not appear to be depleted in the Moon⁸, setting an upper limit on the temperature experienced by proto-lunar material.

If the impactor formed in the region of the terrestrial planets, then it must have been somewhat depleted in volatile elements, as there was an earlier, if less dramatic, depletion of these elements in the inner solar nebula. This is shown by the low volatile/refractory element ratios (for example, $K/U = (1-2) \times 10^4$) in the Earth, Venus and Mars, compared with initial solar nebula values of 6×10^4 .

This depletion event must have occurred very soon after the formation of the Solar System ($T_0 = 4.56 \text{ Gyr}$), as shown by the low initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratios in meteorites^{38,39}, which indicate a very early separation of volatile rubidium from refractory strontium. Also, the lunar initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratio is close to that of the basaltic achondrites—meteorites which formed as igneous rocks 4.54 Gyr ago^{38,39}; thus, lunar ^{87}Sr must have evolved in a low-Rb/Sr environment from times close to T_0 .

Accordingly, the Rb/Sr ratio was probably lower than chondritic in the lunar precursor material. As the Earth, the Moon and the CO, CM and CV carbonaceous chondrites display similar V-Cr-Mn patterns (see above), this widespread volatility-related depletion may have been caused by early volatile depletion in the nebula, rather than the giant impact. It seems safe to conclude that the impactor mantle was not composed of material with the composition of CI chondrites, but had undergone a volatile depletion similar to that seen elsewhere in the inner Solar System^{8,37}, and thus had Rb/Sr ratios, V-Cr-Mn patterns and K/U ratios similar to those of the inner planets; additional potassium, other volatile elements and water were lost in the collision. The physics of volatile depletion during lunar formation needs further study²⁵.

The bulk Moon is probably enriched in refractory elements (such as Al, Ti, Ca and U) by a factor of 1.5 compared with the terrestrial mantle, or about 2.5 times the primitive abundances in CI meteorites. The geochemical arguments for this⁸ have been buttressed by geophysical studies, which support a high-alumina Moon⁴⁰. The most recent refinement of the geophysical data⁴¹ concludes that "only in the case of extreme assumptions can critical aspects of bulk lunar composition be demonstrated to be equivalent to the present-day terrestrial mantle: specifically the Moon has an $Mg\# [= \text{Mg}/(\text{Mg} + \text{Fe})]$ that is too low and an alumina abundance that is too high".

Whether the mantle of the impactor was enriched in refractory elements is less certain; alternatively, fractional condensation from a vapour phase could have enriched the Moon in alumina, uranium and the other refractory elements⁸. But such a process might alter isotopic ratios, and no such effects have been observed in the potassium isotopes²⁹. We conclude that the impactor was probably enriched in refractory elements, noting that the Earth is enriched over CI abundances by a factor of 1.5. This raises broad questions about the relative compositions of the terrestrial planets, selective accretion from already differentiated planetesimals, and the effects of massive collisions that remain for future research.

Siderophile elements. The depletion patterns of siderophile elements in the Earth and Moon (Fig. 2) provide constraints on their abundance in the mantle of the impactor and on the size of the impactor's core. Two important characteristics of the

siderophile abundances in the Earth's upper mantle are that the highly siderophile elements (such as rhenium and iridium) seem to be uniformly distributed throughout the upper mantle, and have chondritic relative abundances (Fig. 2). These facts are often attributed to accretion of a late (post-core-formation) veneer of meteoritic material to the terrestrial upper mantle⁴⁵, because the siderophile element concentrations are too high to have been in equilibrium with a metallic core.

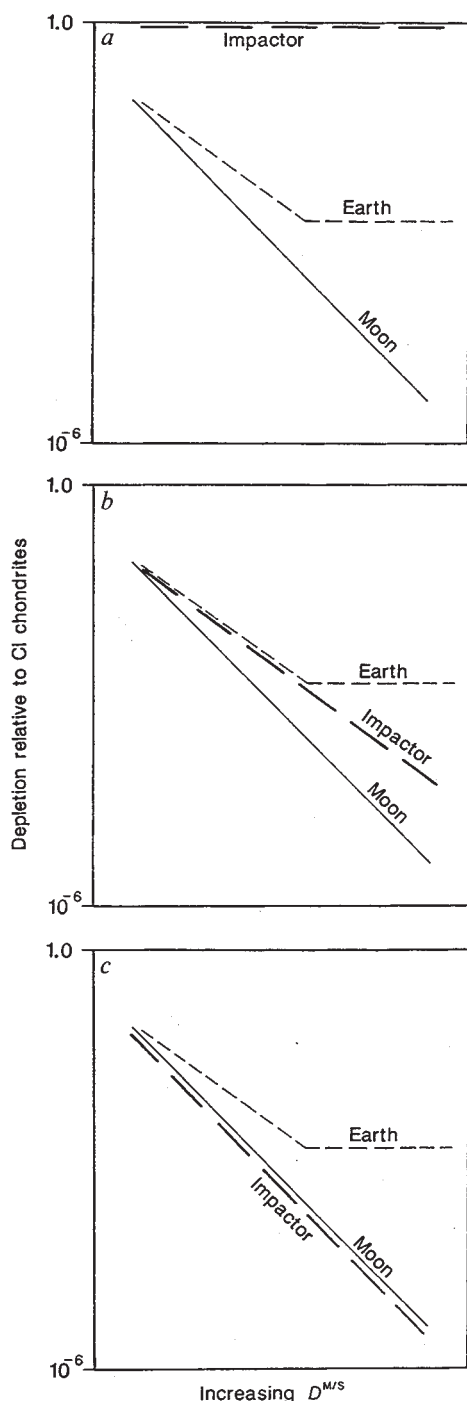


FIG. 3 Schematic siderophile element depletion patterns for the Earth and Moon, resulting from different siderophile depletion patterns in the impactor mantle. *a*, An impactor resembling a chondritic meteorite, with no segregation into metallic core and silicate mantle. *b*, An impactor mantle with a siderophile element pattern similar to that of the Earth's mantle, without the 'late veneer' of highly siderophile elements. *c*, An impactor with a siderophile element depletion pattern similar to that of the lunar mantle.

The lunar siderophile pattern, by contrast, shows a relatively uniform decrease in abundance (relative to chondrites) of the refractory siderophile elements as a function of their siderophile nature. (Elements such as gallium and phosphorus may also be depleted owing to their volatility.) This depletion pattern is quantitatively consistent with equilibrium between the lunar silicates and Fe-rich metal at low degrees of partial melting⁴², presumably during core formation. (The existence of a lunar core of 350–500 km radius, although not conclusively proven, is consistent with many geophysical data, including the Moon's moment of inertia coefficient⁴³, electrical conductivity, and the latest assessment of lunar palaeomagnetism, which seems to require a dynamo in an electrically conducting core between ~4 and 3.5 Gyr ago⁴⁴.) Only the least siderophile elements, such as tungsten and cobalt, are similarly depleted in the Earth and Moon.

The conclusion that core formation can explain the lunar siderophile abundances depends on the lunar core having a relatively low nickel content (<10 wt%⁴²). A nickel-rich core would be less efficient at scavenging siderophiles from the silicate phase; thus, if the lunar core were in fact to contain 41 wt% Ni, as has recently been suggested⁴⁶, the Moon's siderophile depletion would have to be explained by its formation largely out of material from the Earth's (already depleted) mantle. We do not feel that this conclusion is required, however, because the figure of 41% was derived from the Ni abundance in Apollo 15 green glass, assuming⁴⁶ that this glass represents a liquid from the primitive lunar mantle. In fact, the chemistry of the glass—notably its depletion in europium and strontium, characteristic of other lunar mare basalt samples²⁷—indicates that its source region was formed by crystallization from the lunar magma ocean. If this magma ocean was in direct equilibrium with the lunar core, then the Ni content of the green glass would indeed reflect that of the lunar core. This is unlikely, however, because the metal separation would have occurred at some moderate degree of partial melting, whereas the magma ocean probably represents a high degree of melting of the lunar mantle. Assuming that a large proportion of the Moon (or precursor planet) experienced metal segregation at 10% partial melting, following which the silicates were partially or completely melted to form a magma ocean, the nickel content of the metal in the lunar core could range from 10 to 27 wt%.

An additional argument against Ni-rich metal in equilibrium with lunar silicates comes from the very reduced oxidation state of the Moon, which is characterized by a total lack of ferric iron. Lunar basalts crystallized at oxygen fugacities four to five orders of magnitude lower than those of terrestrial basalts at equivalent temperatures⁴⁷. If lunar silicates had equilibrated with Ni-rich metal, lunar oxygen fugacities would be much higher.

The lunar siderophile element pattern allows four possibilities for the siderophile abundances in the impactor, the implications of which are discussed next. The second and third cases seem most plausible.

An undifferentiated impactor. The least likely possibility is that the impactor had chondritic abundances of siderophile elements in its mantle; that is, there was no separation of metal from silicate (Fig. 3*a*). Not only would the lunar iron content derived from such a source be too high, but it is hard to see how a planet larger than Mars could escape differentiation into a separate core and mantle⁴⁸. There is abundant evidence from over 60 varieties of iron meteorite that core formation occurs even in relatively small asteroidal bodies. Also, the absence of a core in the impactor would make it more difficult to place a lunar mass of material in orbit^{20–22}. The objections to an undifferentiated chondritic impactor, and to an oxidized chondritic impactor containing no Fe–Ni metal, are discussed in more detail in ref. 49.

In this model the entire siderophile element depletion observed in lunar silicates has to be explained by core formation

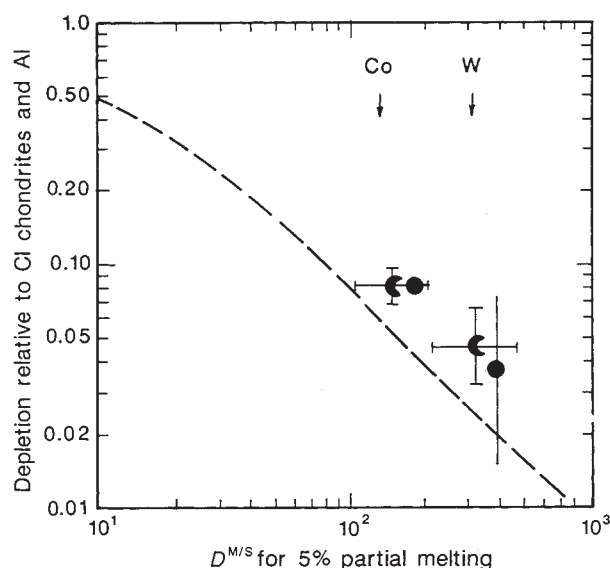


FIG. 4 The maximum core size in the impactor is constrained by the similar abundances of W and Co in the Earth (●) and Moon (○). These imply that the impactor also had similar concentrations of W and Co. The maximum allowable core size (10 wt%) corresponds to the minimum metal/silicate partition coefficients ($D^{M/S}$) for these elements, and the maximum depletion⁴². The minimum partition coefficients are calculated for a composition of 50 wt% S-rich metal and 50 wt% Fe-metal^{42,54}; a larger amount of S-rich metal would reduce the partition coefficient of W too much relative to Co. The maximum depletion for these elements is derived from estimates of the uncertainties of the depletions⁴². The dashed line represents the depletions expected for a core size of 10 wt%, and this line just grazes the maximum depletions observed in the Earth and Moon.

in the Moon. Newsom⁴² concluded that the observed lunar depletion could be explained by segregation of ~4–5.5 wt% Fe–Ni metal^{40,41,50}, assuming segregation at low degrees of partial melting of the silicates (5–9%).

Siderophile depletion in impactor mantle similar to Earth's mantle. With the exception of the highly siderophile elements (Fig. 3b), this could be produced by segregation of a small (<10 wt%) core, assuming equilibrium between the impactor's core and mantle. In the absence of core–mantle equilibrium, as in the Earth, any size of core would be possible, including 30 wt%^{20–22}. If the Moon formed from material with a siderophile element pattern characteristic of the Earth's mantle, including Earth-like abundances of highly siderophile elements, then segregation of additional metal (<0.2 wt%) is needed in the Moon to produce the observed lunar pattern⁴².

Siderophile depletion in impactor mantle similar to present Moon. Formation of the Moon from material with essentially its present abundance of siderophile elements constrains the amount of metal in the impactor that could produce the depletion. Assuming impactor core–mantle equilibrium and the largest possible depletion of Co and W in lunar silicates, the maximum size of the impactor core is ~10 wt% metal (Fig. 4). The composition of the impactor core is also constrained by the Co and W abundances. The metal/silicate partition coefficients⁵¹ are lower for sulphur-rich metal (25–30% S), allowing larger core sizes than for sulphur-free metal (Fig. 3c). Assuming that the impactor core consists of both S-rich and S-poor metal, the fraction of S-rich metal in the core must be <50%. A larger amount of S-rich metal would reduce the partition coefficient for W too much relative to Co, and W would not have been depleted, as is observed for the shergotite parent body^{52,53}. In this model, no metal segregation in the Moon is required to explain the siderophile abundances. If a lunar core exists, as indicated by the geophysical data, there must have been disequilibrium

between the core and the lunar mantle during core formation. *Siderophile depletion in impactor mantle greater than in the Moon.* This possibility is unlikely, as it requires siderophile elements to be added to the Moon from elsewhere. The moderately siderophile elements Co and W have similar abundances in the Earth's mantle and the silicate portion of the Moon (Fig. 4). If the impactor had W and Co abundances significantly below the lunar (and terrestrial) level, addition of terrestrial material would not bring the abundance up to the lunar level. The abundances of the highly siderophile elements could be consistent with the addition of a relatively large amount of terrestrial mantle material to the Moon (15–50%), although additional complications arise if the Earth's mantle already possessed its veneer of highly siderophile elements.

Implications for the Earth

The implications for the Earth of the single impact origin of the Moon are considerable. The event probably triggered or enhanced complete melting of the Earth's mantle²⁶. In addition to the accretion of the impactor's core, the impactor's mantle would have provided about 10% of the mass of the Earth's mantle.

For the second and third models above, with terrestrial or lunar siderophile abundances in the impactor, the implications for the siderophile element budget of the Earth depend on the fate of the metal core from the impactor. Models of the collision^{20–22} indicate that most of the impactor core ends up in the Earth, with the metal penetrating the mantle and wrapping around the Earth's core. This would not disturb the siderophile abundance patterns already present in the Earth's mantle, but a significant amount of material from the impactor's core, enriched in siderophile elements, would probably have been vaporized and redistributed into the mantle. The detailed implications for terrestrial siderophile abundances depend on the fraction of the accreting metal core with small enough grain size to equilibrate with the mantle, but partial retention of metal from the impactor could explain the entire siderophile element pattern in the Earth's mantle^{49,54}. Another variation on this theme states that the present abundances of highly siderophile elements in the Earth's mantle (the late veneer) result from the addition of a small portion of the impactor's core, while the rest of the impactor's core accreted to the Earth's core without significant interaction with the Earth's mantle. The Moon shows little evidence of a late veneer of siderophile elements, supporting an event unique to the Earth as the explanation for the abundances of the highly siderophile elements in the Earth's mantle.

To obtain the observed mantle abundances of the highly siderophile elements requires adding material with chondritic abundances equal to 0.74% of the mass of the Earth's mantle⁵⁵. This can be achieved with 3–4% of the impactor's core, assuming that the total mass of the impactor is between 0.12 and 0.17 Earth masses, respectively. The percentage of the impactor core required depends on the size of the impactor, and is independent of the size of the impactor core because the highly siderophile elements from the impactor would be quantitatively concentrated in the metal. The amount of metal containing the siderophile elements added to the mantle is very small: 0.2 wt% of the Earth's mantle for an impactor core size of 31 wt%, and less than that for smaller cores.

As for the moderate siderophiles, the probable differentiation of the impactor into mantle and core would have depleted these elements in the impactor mantle to levels at or below terrestrial mantle values. Thus the addition of impactor mantle amounting to <10% of the mass of the terrestrial mantle would not significantly alter the terrestrial abundances of the moderately siderophile elements. The contribution of both moderately and highly siderophile elements from the impactor's core to the abundances in the Earth mantle would be only ~1%, insignificant compared with the 5–10% levels in the Earth's mantle.

Conclusions

From these arguments we deduce that not only the dynamics, but the chemical properties of the Earth-Moon system can be explained by a low-velocity (5 km s^{-1}) collision with the Earth of a body of ~ 0.14 Earth masses (larger than Mars) during the final stages of the accretion of the Earth from a hierarchy of planetesimals. The impacting body was already differentiated into a metallic core and silicate mantle; the material that ended up in the Moon was derived principally from the impactor's mantle, which must have been depleted in siderophile elements by core segregation, leaving siderophile abundances somewhere between lunar and present terrestrial mantle values. The core of the impactor was $<10\%$ of its mass, if there was equilibrium between it and the mantle, and the fraction of S-rich metal in the impactor's core was $<50\%$ of the core's mass. The core might have been larger, if the core and mantle were not in equilibrium. The impactor's mantle had volatile-element abundances similar to those of the terrestrial planets, and was enriched in FeO compared to the Earth.

The fate of the impactor's core has important implications for the abundances of siderophile elements in the Earth's mantle. The chondritic relative abundances of the highly siderophile elements can be explained if $\sim 3\text{--}4\%$ of the impactor's core were mixed into the Earth's mantle, the remainder of the core accreting to the Earth's core without interacting with the mantle.

For fuller substantiation, the giant impact hypothesis requires more data, some of which can be collected only on future space missions. For example, the geochemical behaviour of the elements under the extreme conditions of the impact is poorly

understood and it is not clear how much of the depletion of volatile elements and enrichment of refractory elements in the Moon is a result of the impact. To answer this, we need more information on the overall composition of the lunar surface, which could be obtained by an orbiting probe; additional data on heat flow, which would put constraints on bulk composition; and seismic measurements to determine the size of the lunar core and confirm the density discontinuity at 500 km depth. Future lunar missions might obtain direct samples of the lunar mantle from the central peak of Copernicus where mantle rock has evidently been exposed by rebound following the impact that formed the crater.

On Earth, we need to know more about the chemistry of siderophile elements in the mantle, and more about the lower mantle, which remains largely unexplored. Further from home, mechanisms of planetary accretion have to be examined. This will require an understanding of the variations in the composition of the original nebula, which can be estimated from the geochemistry of Mercury, Venus and Mars. The forthcoming exploration of the martian moon, Phobos, and probes to distinct regions of the asteroid belt may also provide insights into the formation of the now-vanished hierarchy of planetesimals that accreted to form the terrestrial planets, and of the one that provided us with our unique and poetically inspiring Moon. □

Horton E. Newsom is at the Institute of Meteoritics and Department of Geology, University of New Mexico, Albuquerque, New Mexico 87131, USA. Stuart Ross Taylor is at the Research School of Earth Sciences, The Australian National University, Canberra, ACT 2601, Australia, and Lunar and Planetary Institute, Houston, Texas 77058-4399, USA.

1. Boss, A. P. & Peale, S. J. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 59–101 (Lunar and Planetary Institute, Houston, 1986).
2. Urey, H. C. in *Physics and Astronomy of the Moon* (ed. Kopal, Z.) 481–523 (Academic, New York, 1962).
3. Ruskol, E. L. *NASA spec. Publ.* **370**, 815–822 (1972).
4. Mizuno, H. & Boss, A. P. *Icarus* **63**, 109–133 (1984).
5. Herbert, F., Davis, D. R. & Weidenschilling, S. J. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 701–730 (Lunar and Planetary Institute, Houston, 1986).
6. Weidenschilling, S. J. et al. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 731–762 (Lunar and Planetary Institute, Houston, 1986).
7. Darwin, G. H. *Phil. Trans. R. Soc.* **170**, 442–530 (1879).
8. Taylor, S. R. *Geochim. cosmochim. Acta* **51**, 1297–1310 (1987).
9. Ringwood, A. E. *Nature* **322**, 323–328 (1986).
10. Ringwood, A. E. & Seifert, S. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 249–278 (Lunar and Planetary Institute, Houston, 1986).
11. Cameron, A. G. W. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 609–620 (Lunar and Planetary Institute, Houston, 1986).
12. Wetherill, G. W. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 519–550 (Lunar and Planetary Institute, Houston, 1986).
13. Mizuno, H., Nakazawa, K. & Hayashi, C. *Earth planet. Sci. Lett.* **50**, 202–210 (1980).
14. Cameron, A. G. W. & Ward, W. R. *Lunar Sci.* **7**, 120–122 (1976).
15. Hartmann, W. K. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 579–608 (Lunar and Planetary Institute, Houston, 1986).
16. Wilhelms, D. E. *Prof. Pap. U.S. geol. Surv.* No. 1348 (1987).
17. Cameron, A. G. W. & Benz, W. *Lunar planet. Sci.* **18**, 151 (1987).
18. Fegley, B. & Cameron, A. G. W. *Earth planet. Sci. Lett.* **82**, 207–222 (1987).
19. Potter, A. E. & Morgan, T. H. *Science* **229**, 651–653 (1985).
20. Benz, W., Slattery, W. L. & Cameron, A. G. W. *Icarus* **66**, 515–535 (1986).
21. Benz, W., Slattery, W. L. & Cameron, A. G. W. *Icarus* **71**, 30–45 (1987).
22. Benz, W., Cameron, A. G. W. & Melosh, H. J. *Lunar planet. Sci.* **19**, 61–62 (1988).
23. Melosh, H. J. & Sonett, C. P. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 621–642 (Lunar and Planetary Institute, Houston, 1986).
24. Kipp, M. E. & Melosh, H. J. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 643–647 (Lunar and Planetary Institute, Houston, 1986).
25. Stevenson, D. J. *A. Rev. Earth planet. Sci.* **15**, 271–315 (1987).
26. Stevenson, D. J. *Lunar planet. Sci.* **19**, 1125–1126 (1988).
27. Taylor, S. R. *Planetary Science: A Lunar Perspective* (Lunar and Planetary Institute, Houston, 1982).
28. Drake, M. J. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 105–124 (Lunar and Planetary Institute, Houston, 1986).
29. Hinton, R. W., Clayton, R. N., Davis, A. M. & Olsen, E. J. *Lunar planet. Sci.* **19**, 497–498 (1988).
30. Wänke, H. & Dreibus, G. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 649–672 (Lunar and Planetary Institute, Houston, 1986).
31. Drake, M. J. et al. *Eos* **69**, 393 (1988).
32. Newsom, H. E. & Drake, M. J. *Lunar planet. Sci.* **18**, 716–717 (1987).
33. Rammensee, W., Palme, H. & Wänke, H. *Lunar planet. Sci.* **14**, 628–629 (1983).
34. Brey, G. & Wänke, H. *Lunar planet. Sci.* **14**, 71–72 (1983).
35. Klock, W. & Palme, H. *proc. lunar planet. Sci. Conf.* **18**, 471–483 (1988).
36. Seifert, S. & Ringwood, A. E. *Earth Moon Planets* **40**, 45–70 (1988).
37. Taylor, S. R. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Matthews, M. S.) 512–534 (University of Arizona Press, Tucson, 1988).
38. Tilton, G. R. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Matthews, M. S.) 259–275 (University of Arizona Press, Tucson, 1988).
39. Brannon, J. C., Podosek, F. A. & Lugmair, G. W. *Proc. lunar planet. Sci. Conf.* **18**, 555–564 (1988).
40. Hood, L. L. & Jones, J. H. *Lunar planet. Sci.* **17**, 354–355 (1986).
41. Mueller, S., Taylor, G. J. & Phillips, R. J. *J. geophys. Res.* **93**, 6338–6352 (1988).
42. Newsom, H. E. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 203–229 (Lunar and Planetary Institute, Houston, 1986).
43. Ferrari, A. J., Sinclair, W. S., Sjogren, W. L., Williams, J. G. & Yoder, C. F. *J. geophys. Res.* **85**, 3939–3951 (1980).
44. Fuller, M. & Cisowski, S. M. *Geomagnetism* **2**, 307–455 (1987).
45. Ganapathy, R. & Anders, E. *Proc. lunar Sci. Conf.* **5**, 1181–1206 (1974).
46. Seifert, S., O'Neill, H. St C. & Brey, G. *Geochim. cosmochim. Acta* **52**, 603–616 (1988).
47. Basaltic Volcanism Study Project *Basaltic volcanism on the Terrestrial Planets* (Pergamon, New York, 1981).
48. Stevenson, D. J. *Science* **214**, 611–619 (1981).
49. Drake, M. J. *J. geophys. Res.* **92**, E377–E386 (1987).
50. Hood, L. L. in *Origin of the Moon* (eds Hartmann, W. K., Phillips, R. J. & Taylor, G. J.) 361–410 (Lunar and Planetary Institute, Houston, 1986).
51. Newsom, H. E. & Palme, H. *Earth planet. Sci. Lett.* **69**, 354–364 (1984).
52. Dreibus, G. & Wänke, H. *Meteoritics* **20**, 367–381 (1985).
53. Treiman, A. *J. geophys. Res.* **92**, E627–E632 (1987).
54. Jones, J. H. & Drake, M. J. *Geochim. cosmochim. Acta* **47**, 1199–1209 (1983).
55. Chou, C.-L., Shaw, D. M. & Crocket, J. H. *J. geophys. Res.* **88**, A507–A518 (1985).
56. Anders, E. & Grevesse, N. *Geochim. cosmochim. Acta* (in the press).

ACKNOWLEDGEMENTS. We thank J. H. Jones, G. J. Taylor, K. Keil, E. R. D. Scott and H. Wänke for useful discussions, and J. H. Jones, M. J. Drake and H. Palme for constructive reviews. H.N. was supported by NASA (K. Keil, principal investigator) and NSF (H.N., principal investigator).

An initiation site for meiotic gene conversion in the yeast *Saccharomyces cerevisiae*

Alain Nicolas*, Douglas Treco, Neil P. Schultes & Jack W. Szostak

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

An initiation site for meiotic gene conversion has been identified in the promoter region of the *ARG4* gene of *Saccharomyces cerevisiae*. The chromosome on which initiation occurs is the recipient of genetic information during gene conversion.

CONVENTIONAL genetic analysis has failed to provide a clear understanding of the mechanism of homologous recombination in meiosis. The mechanism for initiation of recombination remains particularly controversial, as most of the genetic data can be explained by models that invoke either double-strand breaks¹ or single-strand nicks² as the initiating DNA lesion. The existence of specific sites for the initiation of recombination has been suggested³ to explain the genetic phenomenon of polarity: a gradient in the frequency of gene conversion from one end of a gene to the other (reviewed in refs 4 and 5). No such site has been identified and characterized however. Here, and in the accompanying paper⁶, we describe the localization and characterization of a recombination initiation site near the *ARG4* gene of the yeast *S. cerevisiae*. The properties of this site provide evidence for the double-strand-break repair model¹ of meiotic recombination.

We have used classical genetic analysis of deletions constructed by recombinant DNA methods to map the *ARG4* initiation site. A diploid cell that is heterozygous at a mutant site (*m*/+) usually produces two *m* (mutant) spores and two + (wild-type) spores (a mendelian segregation). In some fraction of meioses, however, gene conversion events give rise to progeny where the ratio of + to *m* is 3:1 or 1:3 (non-mendelian segregation). In these cases the information at the position of the mutation on one chromatid of a pair of sister chromatids is lost and replaced with the corresponding information from its homologue. To search for an initiation site, we first re-examined the pattern of recombination within *ARG4*, followed by a deletion analysis to identify sequences essential for high levels of meiotic gene conversion within this gene.

Polarity of gene conversion within *ARG4*

Meiotic recombination in the *ARG4* gene has been studied extensively by Fogel and colleagues (reviewed in ref. 7). The frequency at which an *arg4*⁻ allele engages in gene conversion depends on its position, with the frequency diminishing from one end of the gene to the other. To eliminate the possibility that the variation in conversion frequency was a result of marker effects, silent base substitutions or strain background effects, we re-examined the pattern of recombination within *ARG4* using a set of alleles constructed *in vitro*. Seven mutant alleles were constructed from cloned *ARG4* DNA and transplanted into the yeast genome at the *ARG4* locus, without the addition of any other genetic material. All seven mutations destroy recognition sites for restriction enzymes. Six are insertions or deletions of two to four base pairs (bp) (Fig. 1) and one is a G→C mutation at position +3 of the coding region.

Strains carrying each of these *arg4*⁻ alleles were crossed with wild-type (*ARG4*⁺) haploids, and tetrads were dissected to determine the gene conversion frequency of each allele (Fig. 1). There is a clear gradient in gene conversion frequencies, with

5' alleles having high frequencies and 3' alleles having much lower ones. The values range from 9.1% to 0.4% of total meioses. The gradient of conversion is seen even for mutations of the same physical nature. The *Acc* and *Aha* mutations, both of which are 2-bp additions, have a fourfold difference in conversion frequency. Likewise, the *Bcl* and *Bgl* mutations, both 4-bp additions, convert at frequencies that differ by a factor of eight.

Although all the addition and deletion mutants showed only 3⁺:1⁻ and 1⁺:3⁻ aberrant tetrad classes, the G→C mutation was characterized by a relatively high frequency of tetrads undergoing post-meiotic segregation (reviewed in ref. 4), amounting to about 20% of all aberrant segregations.

The pattern of meiotic recombination at *ARG4* was also examined in a diploid heteroallelic for the *RV* and *Bgl* mutations, which are separated by 1,015 nucleotides. Gene conversion accounted for 88% of the intragenic recombination events, and 83% of these were single-site events that included only the *RV* site. These data are consistent with the polarity observed in single-point crosses, and allowed us to consider the frequency of gene conversion at the *RV* site as being representative of the level of genetic recombination within *ARG4*.

These studies, using *in vitro* constructed mutations, confirm the original observations of polarity within *ARG4*⁷, and show that this phenomenon is independent of the particular genetic background and alleles used.

Homozygous deletions

To identify DNA sequences that are important for gene conversion within *ARG4* we used deletion analysis of a 9-kilobase (kb) region that extends from 7.5 kb upstream of the gene to a position 172 bp 3' to the *ARG4* coding region. We expected three possible deletion phenotypes, assuming that a single site was responsible for the gene conversion gradient. First, deletions upstream of the initiation site or downstream of the *ARG4* marker being tested should have no effect on gene conversion. Second, deletions that remove or inactivate the site should lower the conversion frequency within *ARG4*. Finally, deletions that remove DNA between the site and the marker being assayed should increase the conversion frequency of that marker by moving the marker closer to the initiation site.

We have used the frequency at which the *arg4*-*RV* allele converts to monitor the effects of 11 different homozygous deletions. The conversion frequency of this marker is not affected by deletions that lie upstream of position -316 or downstream of position +345 ($\Delta 1$, $\Delta 2$, $\Delta 3$ and $\Delta 15$, Fig. 2a). In contrast, deletions $\Delta 4$, $\Delta 5$, $\Delta 6$, $\Delta 7$, $\Delta 8$ and $\Delta 9$ (Fig. 2b) all reduce the frequency of gene conversion for *arg4*-*RV*. The reductions, which range from two to ninefold, are all significantly lower than the wild-type level of 7.4%. All these deletions have 3' endpoints between positions -316 and +1, a region that includes all of the *ARG4* promoter⁸.

These deletions define a site that is required for the high frequency of gene conversion at the 5' end of the *ARG4* gene. If this site is an initiation site, deletions between this site and

* Present address: Laboratoire I.M.G. Batiment 400, Université Paris-Sud, 91405 Orsay-cedex, France.

a marker should increase the conversion frequency of that marker. We have used the *arg4-Bgl* and *-Dra* mutations to look for such an effect (Fig. 3). Deletion $\Delta 14$ removes 980 bp of DNA from the coding region and moves the *Bgl* marker almost as close to the proposed initiation site as the *RV* marker is in the wild-type gene. In the presence of this homozygous deletion, gene conversion at the *Bgl* site increases from its normal frequency, 0.4%, to 7.2%, roughly the same as that at the *RV* site in non-deletion controls. Two control crosses confirm that sequences mapping to the region defined by the deletions in Fig. 2 are responsible for this effect. One of these extends the deletion by 142 bp further 5' than $\Delta 14$ (deletion $\Delta 13$), and the second extends it 316 bp further 5' than $\Delta 14$ (deletion $\Delta 12$). Gene conversion of *arg4-Bgl* was 1.2% with $\Delta 13$, and 1.0%

ARG4								
	NspH	Acc	RV	Bcl	Dra	Aha	Bgl	
Mutation(bp)	G→C	+2	-2	+4	-3	+2	+4	
Position(bp)	+3	131	260	345	601	830	1,274	
3+:1-	21	22	31	13	11	1	0	
1+:3-	17	20	31	7	10	6	3	
p.m.s.	10	0	0	0	0	0	0	
Total tetrads	530	674	838	584	742	480	676	
Conversion frequency (%)	9.1	6.2	7.4	3.4	2.8	1.5	0.4	

FIG. 1 Gene conversion frequencies for *in vitro* constructed *ARG4* mutations. Seven *arg4⁻* alleles were constructed *in vitro* and transplanted into the *S. cerevisiae* genome at the *ARG4* locus. The mutations denoted *Acc*, *RV*, *Bcl*, *Dra*, *Aha* and *Bgl* are small insertions or deletions that destroy the recognition sites for the restriction enzymes *AccI*, *EcoRV*, *BclI*, *DraIII*, *AhaI* and *BglII* respectively. The G→C mutation at position +3 is a single base substitution that destroys an *NspHI* site. Nucleotide position +1 is the first base pair in the *ARG4* coding region⁸. The number of conversion events in each direction is listed separately, and combined to yield the conversion frequency. Of the ten tetrads showing post-meiotic segregation (p.m.s.) for the mutation at position +3, one was 5⁺:3⁻ and nine were 3⁺:5⁻.

METHODS. Standard methods for manipulating and modifying DNA were used to construct the *ARG4* mutations²⁰. Plasmid p(SPO13)2 was the source of *ARG4* DNA used in these studies²¹. The G→C transversion mutation was generated by oligonucleotide-directed mutagenesis²² using a synthetic 25-mer with a single base difference from the wild-type *ARG4* sequence at the third base of the AUG initiation codon. All of the mutated plasmids were transformed into *Escherichia coli* strain NPS/RK2 (relevant genotype *argH1 leuB6 trpC1117 pyrF dam⁻*). The *argH1* mutation in this strain can be complemented by the wild-type *ARG4* gene²³, but not by any of the seven mutant genes. All yeast strains used were derived by lithium acetate transformation²⁴ of the isogenic strains²⁵ MGA1 (*MAT α leu2-3, 112 ura3-52 his3 Δ 1 trp1-289*), MGA2 (*MAT α ura3-52 his3 Δ 1 trp1-289 ade2*), MGA3 (*MAT α leu2-3, 112 ura3-52 trp1-289*) and MGA11 (*MAT α leu2-3, 112 ura3-52 ade2*). Cycloheximide-resistant derivatives of MGA1 and MGA2 were selected on 10 μ g ml⁻¹ cycloheximide. The *ARG4* mutations were introduced into yeast either by the standard gene-transplacement technique²⁶ or by a co-transformation procedure²⁷. Single *arg4⁻* colonies from each of the seven transformants were crossed to *ARG4⁺* yeast, diploids were selected, sporulated and the resultant asci were dissected to determine the gene conversion frequencies for each *arg4⁻* allele. The segregation of *ARG4* and at least five additional heterozygous markers was examined for each cross, and only tetrads displaying aberrant segregation at a single locus were scored as conversion events. Media and growth conditions were essentially as described²⁸. Diploid cells were pregrown to a cell density of about 8×10^7 cells ml⁻¹ in YPD (1% yeast extract, 2% Bacto-peptone, 2% dextrose), washed once in water and transferred to liquid sporulation medium at a cell density of 2×10^7 cells ml⁻¹. Sporulation was carried out at 30 °C with vigorous shaking in 1% potassium acetate, 0.1% Bacto-yeast extract, 0.05% dextrose plus any nutritional supplement required by the particular diploid. Spore viability typically >95%.

with $\Delta 12$; these values are not significantly different from each other or from the value of 0.4% reported for the *Bgl* site in Fig. 1. These results indicate that the conversion frequency of *arg4-Bgl* can be increased by a factor of at least six by moving it closer to the conversion element (comparing $\Delta 13$ with $\Delta 14$), and that sequences between -139 and +3 are essential for the high frequency of gene conversion. In a similar experiment, we used *arg4-Dra*, which converts at a frequency of 2.8%. Upon deletion of 566 bp of DNA between positions -37 and +529 ($\Delta 11$), the frequency of gene conversion at the *Dra* site increased to 6.2%, an increase of more than twofold over the wild-type frequency at this site, and a value that is not significantly different from the wild-type frequency for *arg4-RV*. The results of these internal deletion studies indicate that most, if not all of the conversion element is 5' to position -37, as both $\Delta 14$ and $\Delta 11$ can raise the frequency at which downstream alleles convert to a level indistinguishable from the frequency determined for *arg4-RV*. These results confirm the notion that the gradient in gene conversion frequency reflects the physical distance between a marker and the initiation site and not marker effects.

The above experiments place the 5' and 3' boundaries of the sequences that stimulate gene conversion between positions -316 and -37. We were therefore surprised that deletion $\Delta 10$ (-316 to +3) did not affect conversion at the *RV* site (Fig. 2b). As ten deletions that remove all or part of this region have significant effects on gene conversion, we suspect that in the $\Delta 10$ deletion, two sequences have been fused together to create a new sequence that can function to initiate gene conversion.

Direction of information transfer

Deletion of an initiation site from one homologue should lead to a bias in the direction of information transfer (disparity) as only the wild-type chromosome can initiate gene conversion. In a cross that is heterozygous for the 142-bp deletion $\Delta 9$ ($\Delta 9$, Fig. 2b; cross D9HET, Fig. 4), there is a fourfold excess of tetrads that are 3 Δ :1+ over those that are 1 Δ :3+, indicating that the wild-type DNA sequence that lies opposite the deletion is the preferred recipient of information in gene conversion. About half (562) of the tetrads reported for this cross were also heterozygous for *arg4-RV*. Six aberrant segregations at this site were observed, and all six were co-converted with the 142-bp heterology. An extra 10 conversion events involved only the deletion and did not include the *RV* site. These observations indicate that conversion events originate in or near the sequences removed by deletion $\Delta 9$.

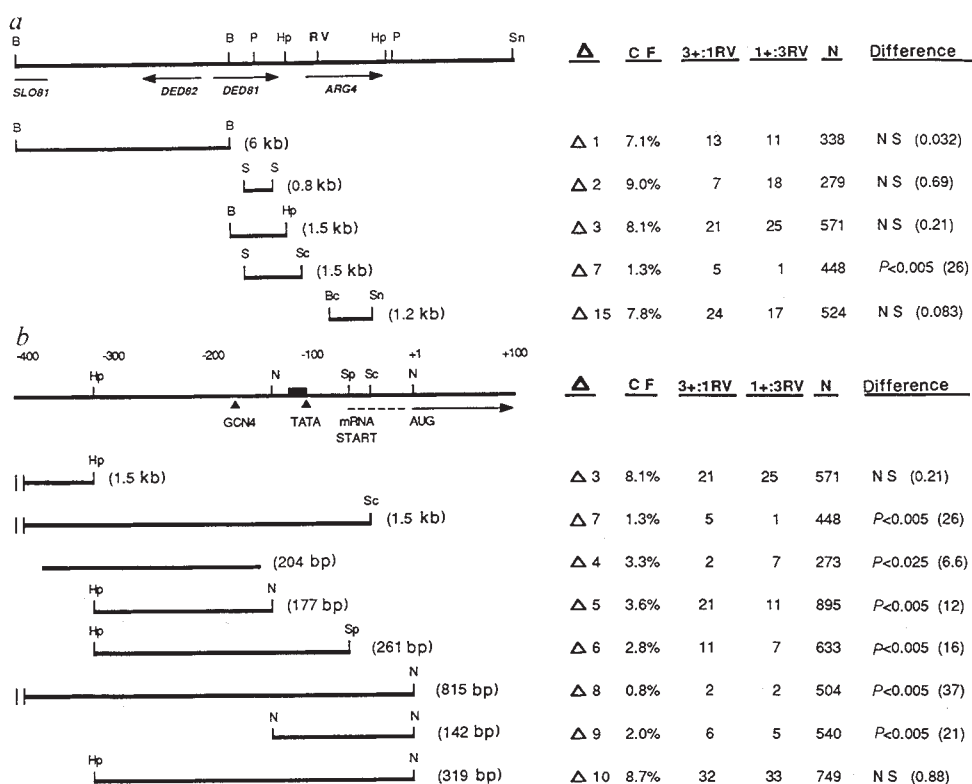
A similar cross between two haploids carrying different deletions results in an 84-bp heterozygosity which includes sequences necessary for the expression of *ARG4* (cross $\Delta 5/\Delta 6$, Fig. 4). There is a 22:1 bias in the direction of information transfer in this cross, with the homologue that retains the sequences from -140 to -56 acting as the preferred recipient in gene conversion. We used two control crosses to show that not all insertion/deletion heterozygosities show this disparity (Fig. 4). One cross had a 137-bp insertion, and the other an 84-bp deletion in the *ARG4* coding sequence. Both of these heterozygosities fall within regions of the gene that are, on the basis of our deletion analysis, downstream of the initiation site. Both heterologies show a modest tendency to convert towards the deletion, or the shorter, of the two recombining *ARG4* regions, but in neither case is the ratio significantly different from an expected ratio of one. It seems that strong disparity is a property only of heterozygous deletions involving the initiation site.

Discussion

Several lines of evidence indicate that a region overlapping the *ARG4* promoter acts as an initiation site for meiotic recombination. First, deletion of the site abolishes the high frequencies of gene conversion seen at the 5' end of the *ARG4* gene. Second, deletions that move the site closer to a marker raise the frequency at which that marker engages in gene conversion. Third,

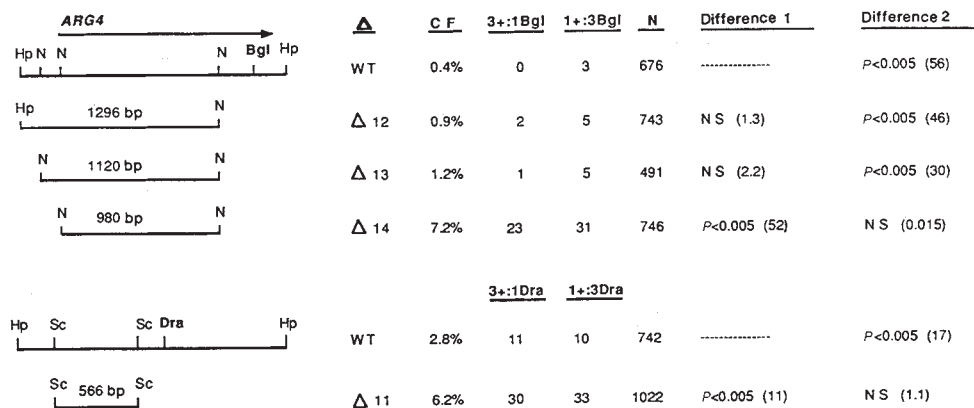
FIG. 2 Effects of homozygous deletions on meiotic gene conversion at the *RV* site of *ARG4*. Solid black bars represent the DNA sequences deleted in each of the 11 crosses assayed. Positions of deletion breakpoints are numbered from the first base in the AUG codon of *ARG4* (position +1). The number of conversion events in each direction is listed separately, and combined to yield the conversion frequency (CF). *N* is the number of four-spore viable tetrads analysed; + indicates wild-type information; *RV* indicates the mutant site. The column labelled 'difference' indicates whether the conversion frequency is significantly different from the value of 7.4% determined for the *RV* site (Fig. 1). The level of significance is indicated by the *P*-value. Also listed are the *G*-values (parentheses) that were used in the significance testing (calculated using the *G*-test of independence²⁹). NS, no significant difference. B, *Bam*HI; Bc, *Bcl*I; Hp, *Hpa*I; N, *Nsp*HI; P, *Pst*I; RV, *Eco*RV; S, *Sal*I; Sc, *Sac*I; Sn, *Sna*BI; Sp, *Ssp*I. a, A 13.5-kb region surrounding the *ARG4* gene and a partial restriction map is shown. Arrows indicate the directions of transcription for the *DED81*, *DED82* and *SLO81* genes. b, A 700-bp region around the 5' end of the *ARG4* gene is shown. The 3' endpoints of $\Delta 3$ and $\Delta 7$ are shown for orientation purposes. The positions of the experimentally determined *GCN4*-binding site³⁰, a possible TATA element, and the measured mRNA start site are indicated. The black box represents a run of 14 A residues (*A_n*). All of these elements have been proposed to be involved in transcription of *ARG4*³¹.

METHODS. Deletions (except $\Delta 4$ and $\Delta 10$) were constructed by cutting with the restriction enzymes indicated. $\Delta 4$ is a spontaneous deletion of 204 bp that arose by selecting for *arg*⁺ function in *E. coli*²³. $\Delta 10$ was constructed by oligonucleotide-directed mutagenesis²² using a synthetic 48-mer that spans the deletion breakpoints. Plasmids containing these deletions were subjected to DNA sequence analysis to ensure that no additional sequences were lost or added. This analysis revealed that $\Delta 8$ extended 55 bp further upstream than the *Sal*I site at position -0.76 kb. Haploid strains were constructed with the *URA3* gene³² inserted at various positions within and upstream of *ARG4*. Deletions were introduced by co-transformation²⁷ with a CEN plasmid followed by screening for loss of the *ura*⁺ phenotype, and then by Southern analysis for the presence of the deletion. Genetic analysis of the *URA3* inserts revealed two essential genes upstream of *ARG4*, *DED81* and *DED82*, and one gene, *SLO81* which results in an extremely slow growth



phenotype. All diploids used in the deletion analysis contain a 10.5 kb insert in the *URA3* gene at its normal chromosomal position³³. This insert, designated *dup*⁺, contains sequences from -7.3 kb to +5.2 kb around *ARG4*, but without a 2-kb fragment from -316 bp to +1,745 bp that carries the *ARG4* gene⁸. The *dup*⁺ construct complements deficiencies created upon disruption of *DED81* or *DED82*. Yeast transformation and tetrad analysis were as described in the Fig. 1 legend. Deletions $\Delta 6$ through $\Delta 10$, and $\Delta 15$ confer an *arg*⁻ phenotype to yeast cells. The *ARG4* genotype of meiotic progeny was determined by allele testing, using a pair of tester strains carrying the *arg4-RV* mutation. Original dissection plates were printed onto a lawn of cells that were *arg4-RV*, *thr4*⁻ and either *MATa* or *MATα*. After mating for 16 h at 30 °C, the plates were printed onto minimal media and irradiated with a 354 nm UV light source for 1 min at distance of 1 m. Cells that are *arg4-RV* cannot give rise to *arg*⁺ papillae by recombination with an *arg4-RV* tester. All tetrads showing non-mendelian segregation were retested with an *arg4-RV* and other *arg4*⁻ testers to ensure that *arg4*⁻ allele was not generated by ectopic recombination with the *dup*⁺ locus at *URA3* on chromosome V. Such events constitute ≤0.5% of all gene conversion events at *ARG4* (data not shown).

FIG. 3 Effects of homozygous deletions on meiotic gene conversion at the *Bgl*I (deletions $\Delta 12$, $\Delta 13$ and $\Delta 14$) and *Dra*I sites ($\Delta 11$) in *ARG4*. Solid black bars represent DNA sequences deleted in each of the experimental crosses. The numbers of conversion events in each direction are listed separately and are combined to yield the conversion frequency. The data for wild-type (WT) levels of conversion at these sites are taken from Fig. 1. The column headed 'difference 1' indicates whether the conversion frequency is significantly different from the wild-type level at that particular site (data from Fig. 1). + indicates wild-type information; *Bgl*I and *Dra*I indicate the mutant sites. The column headed 'difference 2' indicates whether the conversion frequency is significantly different from the wild-type frequency at the *RV* site (data from Fig. 1). Abbreviations and statistical methods are as in Fig. 2. Because deletions $\Delta 11$, $\Delta 12$, $\Delta 13$ and $\Delta 14$ confer an *arg*⁻ phenotype, *ARG4*



genotypes were scored as described in the Fig. 2 legend, using *arg4-Bgl*I or *arg4-Dra*I tester strains. For $\Delta 11$ and $\Delta 14$, the data are pooled from crosses with and without *dup*⁺ homozygous. The *dup*⁺ genotype had no significant effect on the gene conversion frequency (data not shown).

heterozygous deletions show disparity of gene conversion, as expected when initiation occurs preferentially on one chromosome. Fourth, the initiation site is co-converted with the markers upon which it acts, which is consistent with recombination events starting at the initiation site and extending for variable distances away from the site. Finally, as we show in the accompanying paper⁶, we are able to detect the transient appearance of double-strand breaks at the initiation site during meiosis.

The property of disparity is important in considering models for initiation. We find that the chromosome on which events are initiated is the recipient of information in gene conversion. This is consistent with studies on meiotic hotspots in other fungi (for example, the *M26* mutation in the *ADE6* gene of *Schizosaccharomyces pombe*⁹, the *YS17* mutation in the *BUFF* locus of *Sordaria brevicollis*¹⁰, and the *cog*⁺ variant near the *HIS3* gene of *Neurospora crassa*¹¹). It has also recently been shown¹² that a yeast centromere can act to inhibit adjacent sequences from acting as recipients, but not donors, of information during meiotic gene conversion. The simplest explanation for this phenomenon is that the centromere acts in *cis* to inhibit initiation.

It therefore seems to be a general phenomenon that the chromosome on which initiation occurs is the recipient of information in gene conversion. This is inconsistent with the original Aviemore model² for recombination, in which a single-strand displaced from the initiating chromosome is transferred to the recipient chromosome. It is, however, consistent with either the modified Radding model¹³, in which a single-stranded gap on the initiating chromosome is repaired using information from the intact donor chromosome, or with the double-strand-break repair model¹, in which conversion occurs by repair of a double-strand gap on the initiating chromosome using information on the intact donor chromosome.

The inevitable evolutionary consequence of a system in which initiation sites are the preferred recipients in gene conversion is that the sequences will tend to be lost as a result of the event itself. It is paradoxical that sequences that stimulate recombination are lost as a mechanistic consequence of this process, especially as studies on chromosome segregation in meiosis I have shown that recombination is essential for proper chromosome disjunction (reviewed in ref. 14). A possible solution to this paradox is that by coupling the signals for recombination to those for a cellular process that is under strong selective

pressure (for example transcription), any loss of initiation sites for recombination that may occur is compensated for by selection to maintain functional promoters.

The region that we have identified as the initiation site overlaps with the *ARG4* promoter. *ARG4* transcripts increase by 5- to 10-fold within one hour of transfer to sporulation medium (data not shown). Deletions that remove sequences essential for transcription of the *ARG4* gene and give rise to an *arg*⁻ phenotype are highly deficient in *ARG4* transcripts in meiosis and show decreased conversion. With the exception of $\Delta 10$, which we believe creates a cryptic initiation site, no deletion that destroys any part of the promoter is without an effect on gene conversion.

In *S. cerevisiae*, a relationship between transcription and recombination has been demonstrated for mating-type switching¹⁵ and for the *HOT1* site¹⁶, a sequence that stimulates mitotic recombination and consists of the essential sequences for ribosomal DNA transcription. Mitotic recombination between duplicated *GAL10* genes has recently been shown to be strongly stimulated by galactose-induced transcription (B. Thomas and R. Rothstein, personal communication). At present, however, we cannot eliminate the possibility that the *ARG4* promoter and conversion element merely overlap, or that the conversion element is a specific DNA sequence that stimulates recombination in the context of an active promoter.

It is important to note that gene conversion is decreased, but not abolished, in the absence of the initiation site that we have identified. The deletions that have the strongest effects, $\Delta 7$, $\Delta 8$, and $\Delta 9$, still allow 1-2% gene conversion at the *RV* site. This background of gene conversion is probably due to initiation at other sites.

One of the major unanswered questions about meiotic recombination is to what extent heteroduplex DNA contributes to gene conversion. In single-strand initiation models^{2,13}, all conversion occurs by heteroduplex formation followed by mismatch repair. In the double-strand-break repair model¹, gene conversion can occur either by double-strand gap repair or by mismatch repair of heteroduplex DNA. Some heteroduplex DNA is known to form, as post-meiotic segregation is a well-documented phenomenon (reviewed in refs 4 and 5). Studies of post-meiotic segregation suggest that a minimum of 20-30% of all gene conversion events (and perhaps twice as many) between positions +3 and +341 go through a heteroduplex intermediate¹⁷. The suggestion that long regions of heteroduplex DNA can form

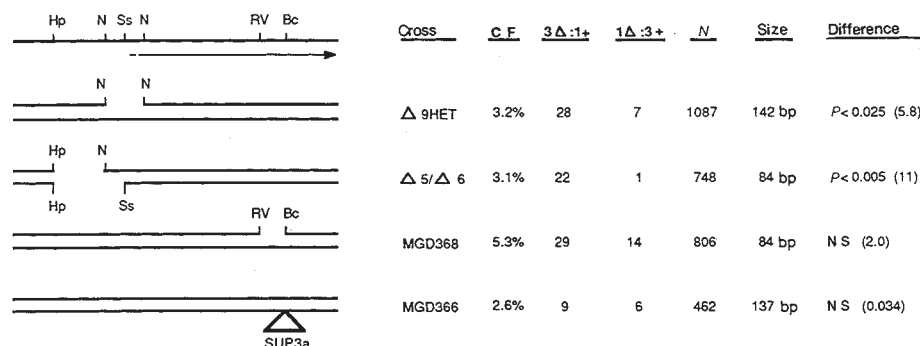


FIG. 4 Gene conversion of heterozygous deletions/insertions at *ARG4*. Data for $\Delta 9$ HET is combined data from two similar crosses (MGD324 and MGD184), one of which (MGD324) was also heterozygous for *arg-RV* in *trans* to the deletion (see text). Both crosses displayed a similar bias in directionality of conversion for the 142 bp deletion (13 3Δ:1⁺ and 3 1Δ:3⁺ in MGD324 and 15 3Δ:1⁺ and 4 1Δ:3⁺ in MGD184). Cross $\Delta 5/\Delta 6$ is described in the text. Conversion events in each direction are listed separately, but are combined to yield the frequency at which the heterology engages in gene conversion. Definitions and abbreviations are as in the Fig. 2 legend. Δ indicates the deletion genotype, + denotes the wild-type configuration. The column headed

'difference' indicates the significance of the deviation from an expected ratio of one for the two conversion classes.

METHODS. Abbreviations and statistical methods are as described in the legend to Fig. 2, except that Yates' correction for small sample size was used in the calculation of G_{adj}^{29} . The 84-bp deletion that is heterozygous in MGD368 was created by digesting *ARG4* DNA with *EcoRV* and *BclI*, filling in the overhanging ends with T4 DNA polymerase and religating. The *SUP3a* fragment used in cross MGD366, a kind gift from Dr R. Rothstein, was inserted as a 137-bp *BamHI* fragment into the *BclI* site of *ARG4*. Yeast co-transformation and genetic analysis was as described in the Fig. 2 legend.

next to a region of double-strand-gap-repair is supported by the finding of extensive single-stranded DNA next to a double-strand break at the *ARG4* initiation site⁶, and by genetic analysis of linear and gapped plasmid DNA after transformation into yeast^{18,19}.

The identification of a naturally occurring initiation site for gene conversion is a first step in the molecular analysis of meiotic

recombination. The further characterization of this sequence within and outside its normal context will contribute to our understanding of how initiation is regulated. By focusing on this DNA sequence throughout meiosis, intermediates in the initiation event can be identified and characterized, allowing us to establish the molecular nature of initiation and distinguish experimentally between the current models for recombination.

Received 2 August 1988; accepted 18 January 1989.

1. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. & Stahl, F. W. *Cell* **33**, 25–35 (1983).
2. Meselson, M. & Radding, C. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 356–361 (1975).
3. Holliday, R. *Genet. Res.* **5**, 282–304 (1964).
4. Whitehouse, H. L. K. *Genetic Recombination: Understanding the Mechanisms* (Wiley, New York, 1982).
5. Orr-Weaver, T. L. & Szostak, J. W. *Microbiol. Rev.* **49**, 33–58 (1985).
6. Sun, H., Treco, D., Schultes, N. & Szostak, J. W. *Nature* (in this issue).
7. Fogel, S., Mortimer, R. K. & Lusnak, K. in *The Molecular Biology of the Yeast Saccharomyces* (eds Strathern, J., Jones, E. & Broach, J.) 289–339 (Cold Spring Harbor, New York, 1981).
8. Beacham, I. R., Schweitzer, B. W., Warwick, H. M. & Carbon, J. *Gene* **29**, 271–279 (1984).
9. Gutz, H. *Genetics* **69**, 317–337 (1971).
10. MacDonald, M. V. & Whitehouse, H. L. K. *Genet. Res.* **34**, 343–380 (1979).
11. Catcheside, D. G. A. *Rev. Genet.* **8**, 279–300 (1974).
12. Lambie, E. J. & Roeder, G. S. *Cell* **52**, 863–873 (1988).
13. Radding, C. M. et al. *Cold Spring Harb. Symp. quant. Biol.* **47**, 821–828 (1982).
14. Murray, A. W. & Szostak, J. W. A. *Rev. Cell. Biol.* **1**, 289–315 (1985).
15. Klar, A. J. S., Strathern, J. N. & Hicks, J. B. *Cell* **25**, 517–524 (1981).
16. Voelkel-Meiman, K., Keil, R. L. & Roeder, G. S. *Cell* **48**, 1071–1079 (1987).
17. White, J. H., Lusnak, K. & Fogel, S. *Nature* **315**, 350–352 (1985).
18. Rothstein, R. *Cold Spring Harb. Symp. quant. Biol.* **49**, 629–637 (1984).
19. Orr-Weaver, T. L., Nicolas, A. & Szostak, J. W. *Molec. cell. Biol.* **8**, 5292–5298 (1988).
20. Ausubel, F. M. et al. *Current Protocols in Molecular Biology* (Greene, New York, 1987).
21. Wang, H.-T., Frackman, S., Kowalsyn, J., Esposito, R. E. & Elder, R. *Molec. cell. Biol.* **7**, 1425–1435 (1987).
22. Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488–492 (1985).
23. Schweitzer, B. thesis, Univ. California, Santa Barbara (1982).
24. Ito, H., Fukuda, Y., Murata, K. & Kimura, A. *J. Bact.* **153**, 163–168 (1983).
25. Treco, D., Thomas, B. & Arnheim, N. *Molec. cell. Biol.* **5**, 2029–2038 (1985).
26. Scherer, S. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4951–4955 (1979).
27. Rudolph, H., Koenig-Rauseo, I. & Hinnen, A. *Gene* **36**, 87–95 (1985).
28. Sherman, F., Fink, G. R. & Lawrence, C. W. *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, New York, 1978).
29. Sokal, R. R. & Rohlf, F. J. *Biometry*, 549–620, (Freeman, San Francisco, 1969).
30. Hope, I. A. & Struhl, K. *Cell* **43**, 177–188 (1985).
31. Struhl, K. *Cell* **49**, 295–297 (1987).
32. Bach, M. L., Lacroute, F. & Botstein, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 386–390 (1979).
33. Mortimer, R. K. & Schild, D. *Microbiol. Rev.* **49**, 181–212 (1985).

ACKNOWLEDGEMENTS. We thank R. Easton Esposito for plasmid p(SP013)2, Andy Ellington, Bernard de Massy and Jean-Luc Rossignol for critical reading of the manuscript, and J. Michael Cherry for oligonucleotide synthesis and for the development of TRAD, a computer program for keyboard entry of tetrad data. This work was supported by Hoechst AG.

Transcription activation by the adenovirus E1a protein

James W. Lillie & Michael R. Green

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

The adenovirus E1a protein stimulates transcription of a wide variety of viral and cellular genes. It is shown here that E1a has the two functions characteristic of a typical cellular activator: one direct E1a to the promoter, perhaps by interacting with a DNA-bound protein, and the other, an activating region, enables the bound activator to stimulate transcription.

RECENT experiments have defined two functional regions of cellular transcription activators^{1–3}. One, the 'binding region', directs the protein to DNA, enabling the other, the 'activating region', to interact with a component of the transcription complex to stimulate transcription. These regions are modular: each is a separable, functional unit that can be joined to the complementary region of an unrelated activator to create a functional hybrid activator¹ (reviewed in ref. 3). In most instances the binding region interacts directly with a specific DNA sequence, but in some cases the binding region recognizes a DNA-bound protein^{4–10}. The binding region determines the activator's specificity by targeting the protein to only those promoters that have the appropriate binding site. The activating regions are short acidic fragments of no specific primary sequence^{11–14}.

The adenovirus E1a protein (E1a) is a powerful transcription activator that is required for the expression of viral early genes^{15,16}. Several lines of evidence suggest that E1a is atypical. First, E1a stimulates transcription of a wide variety of viral and cellular genes that lack a specific promoter element required for induction. Second, E1a enhances expression of genes transcribed by both RNA polymerase II and III (reviewed in ref. 17). Third, E1a does not bind a specific DNA sequence¹⁸. These differences suggest that E1a functions by a novel mechanism. It has been proposed, for example, that E1a stimulates transcription indirectly by converting cellular transcription factors from

an inactive to an active form^{19–22}. Despite these apparent differences we present evidence here that E1a functions like a typical cellular activator.

DNA-bound E1a activates transcription

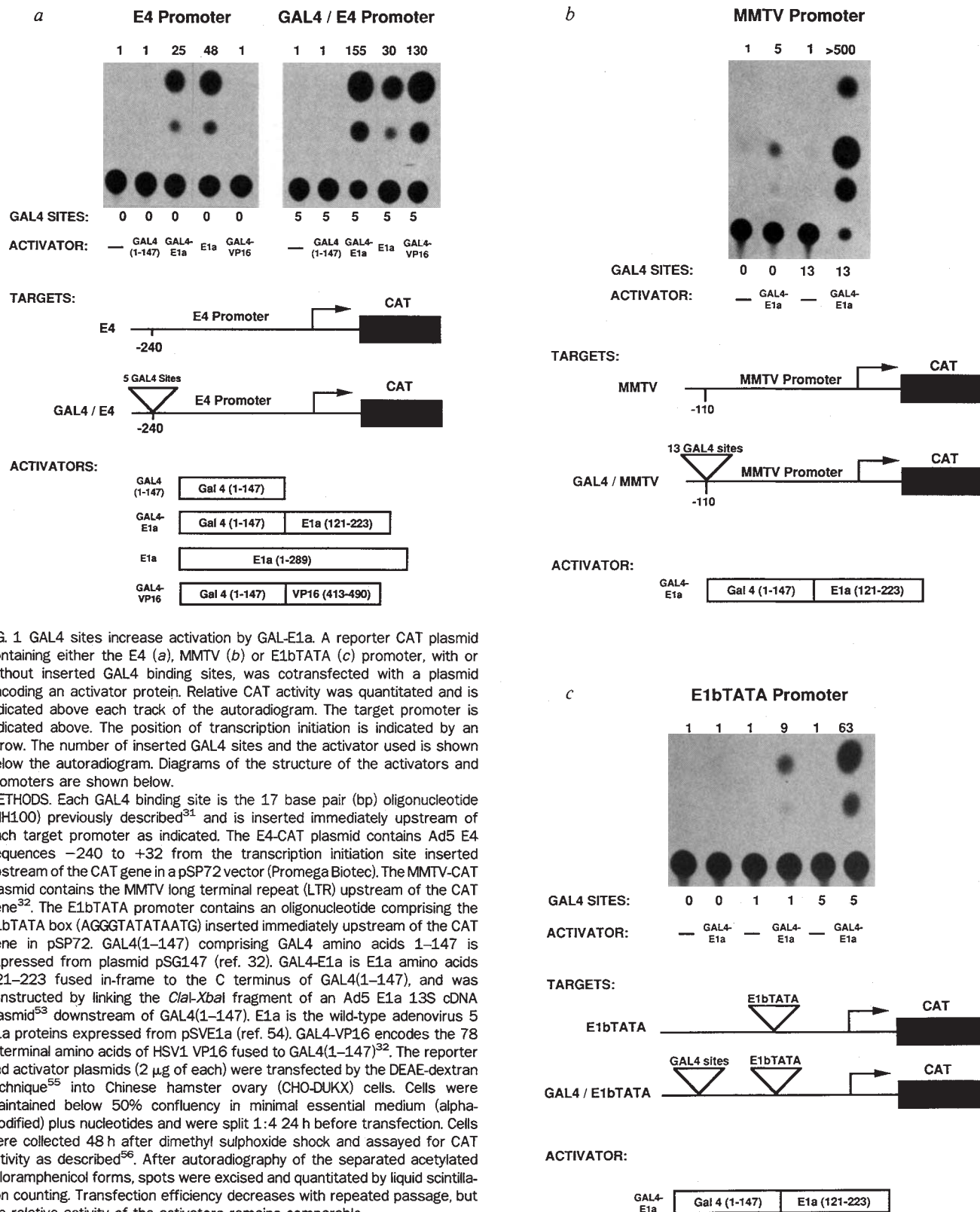
Previous studies have shown that a small, highly conserved E1a region (designated region 3, amino acids 140–189; see ref. 23 and Fig. 6b) is necessary, and in some instances sufficient, to activate transcription^{24–30}. Figure 1 shows that an E1a fragment containing region 3 activates transcription in mammalian cells when bound artificially to the promoter through the DNA-binding region of another activator. For this experiment, we constructed a plasmid that expresses E1a amino acids 121–223 fused to the GAL4 DNA-binding domain, GAL4(1–147). GAL4-E1a contains the GAL4 sequences required to recognize the GAL4-binding site, but lacks the GAL4 sequences that enable promoter-bound GAL4 to activate transcription *in vivo*^{11,12,31–36}. The activity of GAL4-E1a is measured following cotransfection with a 'reporter' plasmid containing the chloramphenicol acetyltransferase (CAT) gene linked to a promoter, with or without inserted GAL4 sites.

Figure 1a shows that GAL4-E1a activates the adenovirus E4 promoter, which lacks GAL4 sites, but does so more efficiently if GAL4 sites are present (Fig. 1a, GAL4-E1a; compare E4 and GAL4/E4 promoters). As expected, wild-type E1a, which lacks

the GAL4 DNA-binding domain, activates the E4 and GAL4/E4 promoters comparably (Fig. 1a, E1a; compare E4 and GAL4/E4 promoters). Binding GAL4(1-147) alone does not enhance activation, because GAL4(1-147) lacks an activating region¹¹ (Fig. 1a).

Insertion of GAL4 sites dramatically increases activation of two other promoters. The mouse mammary tumour virus (MMTV) promoter is activated weakly by GAL4-E1a but the

insertion of GAL4 sites increases activation at least 100-fold (Fig. 1b). We also tested a 'core' promoter, created by inserting a double-stranded oligonucleotide bearing the adenovirus E1b TATA box upstream of the CAT gene. This E1bTATA promoter is not activated detectably by wild-type E1a (data not shown) and activation by GAL4-E1a is detectable only if GAL4 sites are present (Fig. 1c). (We note that weak activation of the E1b TATA box has been observed in virus infection experiments³⁷).



Although activation of the E1bTATA promoter is enhanced by a single GAL4 site, the presence of additional GAL4 sites amplifies the response—presumably by increasing the number of bound GAL4-E1a molecules (Fig. 1c). Taken together, these results indicate that binding E1a to DNA (or binding E1a more efficiently to DNA) enhances its ability to activate transcription. This in turn indicates that E1a may function naturally in the vicinity of the promoter.

E1a contains an activating region

Figures 2, 3 and 4 show that E1a contains a powerful activating region required for the activation of promoters that contain GAL4 sites and those that lack them. Deletion analysis shows that E1a amino acids 139–149 (see Fig. 6c) are essential for the activity of GAL4-E1a. In these experiments, we compare the activity of several N-terminal deletion mutants of the E1a fragment 121–223. As the stability of similar proteins with different

N-terminal sequences varies greatly^{38,39}, and as large variations in protein stability would distort comparisons, each derivative used in these and subsequent experiments begins with the same 147 amino acids, GAL4(1–147), even when the target promoter lacks GAL4 sites.

Removal of E1a amino acids 140–185 (most of region 3) abolishes GAL4-E1a's activity, whether or not GAL4 sites are present (Fig. 2a, GAL4-E1a Δ 3) even when 10 times more DNA is transfected (data not shown). In contrast, deletion of E1a amino acids 121–132 (most of region 2; see ref. 23 and Fig. 6b) does not affect GAL4-E1a's activity on either promoter (Fig. 2b, GAL4-E1a Δ N132). Deletion of the remaining six amino acids of region 2 contiguous to region 3 reduces activation slightly (Fig. 2b, GAL4-E1a Δ N138). Further deletion of the first 10 amino acids of region 3 abolishes GAL4-E1a's activity on both promoters (Fig. 2b, GAL4-E1a Δ N149), even when 10 times more DNA is transfected (data not shown). These results indi-

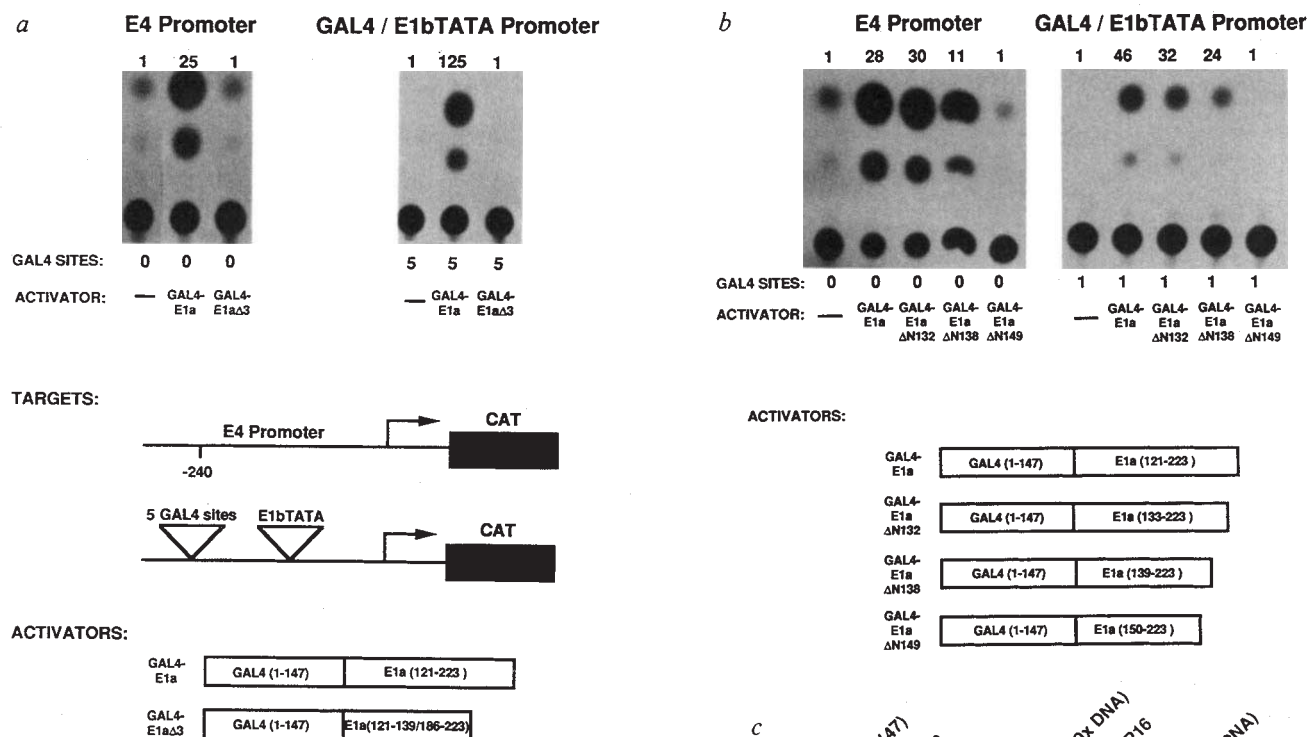


FIG. 2 E1a amino acids 139–149 are essential for activation of promoters with and without GAL4 sites. **a**, **b**, A reporter CAT plasmid containing the E4 or GAL4/E1bTATA promoter was cotransfected with a plasmid encoding the activator indicated below each track of the autoradiogram. Relative CAT activity was quantitated and is indicated above each track. **c**, Protein levels of GAL4-E1a derivatives. ³⁵S-labelled GAL4-E1a derivatives from transfected COS cells were immunoprecipitated with an anti-GAL4(1–147) antibody (a gift of I. Sadowski) and fractionated on a 10% SDS-polyacrylamide gel. The transfected activator is indicated above each lane. Relative molecular mass markers are indicated on the left.

METHODS. Transfections were performed as described in the Fig. 1 legend. Each activator is a derivative of GAL4-E1a. GAL4-E1a Δ 3 lacks the 13S-specific amino acids of region 3, amino acids 140–185 (see Fig. 6b), and was constructed by replacing the *Cla*I–*Xba*I fragment of GAL4-E1a with the *Cla*I–*Xba*I fragment of an Ad5 12S cDNA plasmid⁵³. GAL4-E1a Δ N132 and GAL4-E1a Δ N138 are progressive 5' *Bal*31 deletions of the 13S cDNA *Cla*I–*Xba*I fragment, fused in-frame to GAL4(1–147). Δ N132 begins at Ad5 nucleotide 956, E1a amino acid 133; Δ N138 begins at Ad5 nucleotide 973, amino acid 139. GAL4-E1a Δ N149 contains the *Sma*I–*Xba*I fragment of the 13S cDNA plasmid fused to GAL4(1–147) and begins at Ad5 nucleotide 1010, amino acid 150. COS cells transfected using the DEAE-dextran procedure were labelled with 200 μ Ci ml⁻¹ of [³⁵S]methionine for 3 h, 72 h after the dimethyl sulphoxide shock. Subsequently, whole-cell lysates were cleared with preimmune serum and *Staphylococcus aureus* protein A, the anti-GAL4(1–147) antibody added, and the immune complexes purified and fractionated

on a 10% SDS-polyacrylamide gel. In parallel, unlabelled whole-cell lysates prepared from COS cells cotransfected with a plasmid encoding each activator and the GAL4-E1bTATA reporter CAT plasmid were assayed for CAT activity. The results obtained in COS cells were identical to those obtained in CHO cells (data not shown).

cate that E1a amino acids 139–149 (region 3) are essential for GAL4-E1a's activity on promoters containing or lacking GAL4 sites.

To confirm that the above results reflect differences in activity of the GAL4-E1a derivatives, and not variations in protein concentration, we measured the level of each activator after transfection. Figure 2c shows that the various GAL4-E1a derivatives are expressed comparably and previous studies have shown that proteins of this size will enter the nucleus by passive diffusion^{40–41}.

Figure 3 demonstrates that E1a amino acids 139–149 can be replaced by the acidic activating region of an unrelated activator. The herpes simplex virus 1 (HSV1) VP16 protein is a typical activator in that it contains a promoter-binding function and an acidic activating region⁴². The VP16 activating region is particularly powerful³². Fusion of VP16's activating region to the C terminus of GAL4-E1aΔN149 restores its ability to activate the E4 promoter (Fig. 3, compare GAL4-E1aΔN149 and GAL4-E1aΔN149-VP16). (GAL4-E1aΔN149-VP16 is expressed at a higher level than GAL4-E1aΔN149, Fig. 2c, but this is not sufficient to account for the difference in activity.) Thus, the acidic activating region of the HSV1 VP16 protein can substitute for the natural activity of E1a amino acids 139–149 (as assayed on the E4 promoter), although these two protein fragments bear no obvious sequence similarity. We conclude that this portion of E1a functions solely as an activating region. This experiment is in essence a 'domain swap'^{1,13,32,33,43–45}.

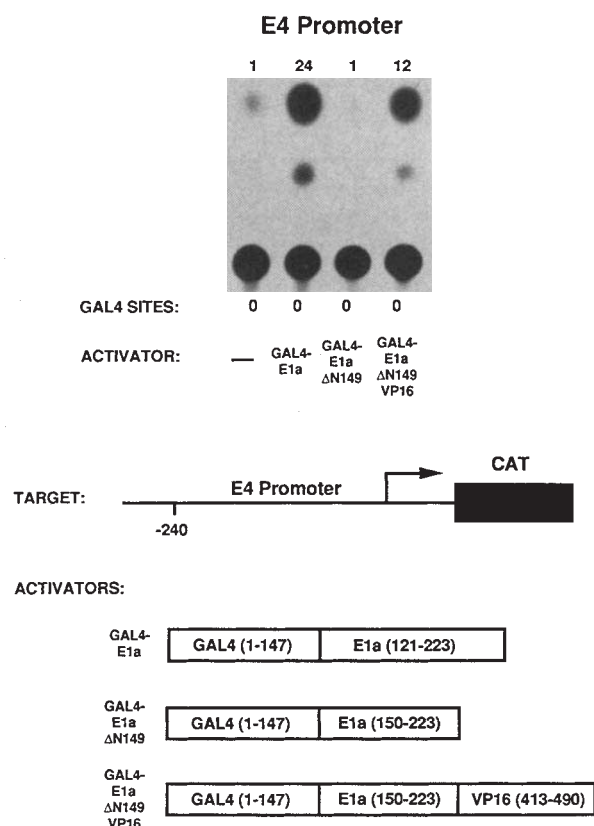


FIG. 3 E1a amino acids 139–149 function as an activating region. A reporter CAT plasmid containing the E4 promoter was cotransfected with a plasmid encoding the activator indicated below each track. Relative CAT activity was quantitated and is indicated above each track.

METHODS. Transfections were performed as described in the Fig. 1 legend. GAL4-E1a and GAL4-E1aΔN149 are described in the legends to Figs 1 and 2, respectively. GAL4-E1aΔN149-VP16 contains the 78 C-terminal amino acids of the HSV1 VP16 protein fused in-frame to the C terminus of GAL4-E1aΔN149.

Finally, GAL4-E1a is a powerful activator, consistent with the notion that E1a, a potent activator, contains a strong activating region. The GAL4/E1bTATA promoter allows us to gauge the strength of E1a's activating region, as transcription is not detectable in the absence of GAL4 sites (Fig. 4). In this assay, E1a's activating region is at least 20 times stronger than the activating regions of either wild-type GAL4 or the GAL4 derivative, GAL4-B17 (Fig. 4, GAL4, GAL4-B17; ref. 12). In fact, the activity of GAL4-E1a is only twofold less than that of GAL4-VP16, which contains the strongest known activating region³² (Fig. 4). The strength of E1a's activating region indicates that its activity is not the adventitious result of fusion to GAL4 (1–147).

E1a contains a promoter-binding region

Like other activating regions, the VP16 activating region must be positioned in the vicinity of the promoter to function (reviewed in ref. 3). Thus, an important implication of the experiment shown in Fig. 3 is that GAL4-E1aΔN149 must contain an activity that directs the VP16 activating region to the wild-type E4 promoter. Neither the GAL4 binding domain, nor the VP16 activating region of GAL4-E1aΔN149-VP16, can provide this activity as GAL4-VP16 fails to activate the E4 promoter

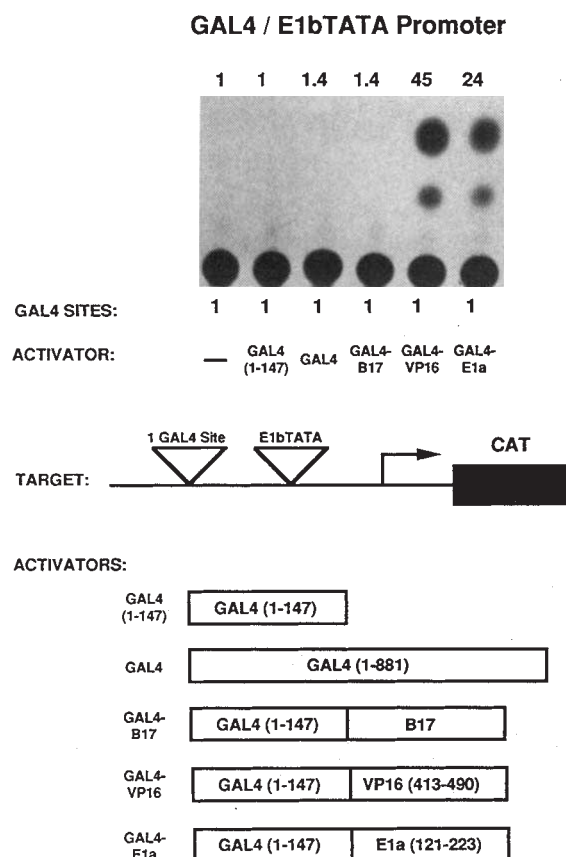


FIG. 4 E1a contains a powerful activating region. A reporter CAT plasmid, containing the E1bTATA promoter with a single GAL4 site inserted immediately upstream, was cotransfected with one of several plasmids that encode an activator. Relative CAT activity was quantitated and is indicated above each track. The activator used is indicated below each track. The E1bTATA promoter is activated detectably by both wild-type GAL4 and GAL4-B17 if an additional four GAL4 sites are inserted upstream, but not as efficiently as it is activated by GAL4-E1a (data not shown).

METHODS. Transfections were as described in the Fig. 1 legend. GAL4(1–147), GAL4-VP16 and GAL4-E1a are as described in Fig. 1. GAL4 is the wild-type GAL4 protein, amino acids 1–881, expressed from pSG4 (ref. 32). GAL4-B17 encodes an acidic *Escherichia coli* sequence fused to GAL4(1–147) and is expressed from pSGB17 (ref. 32).

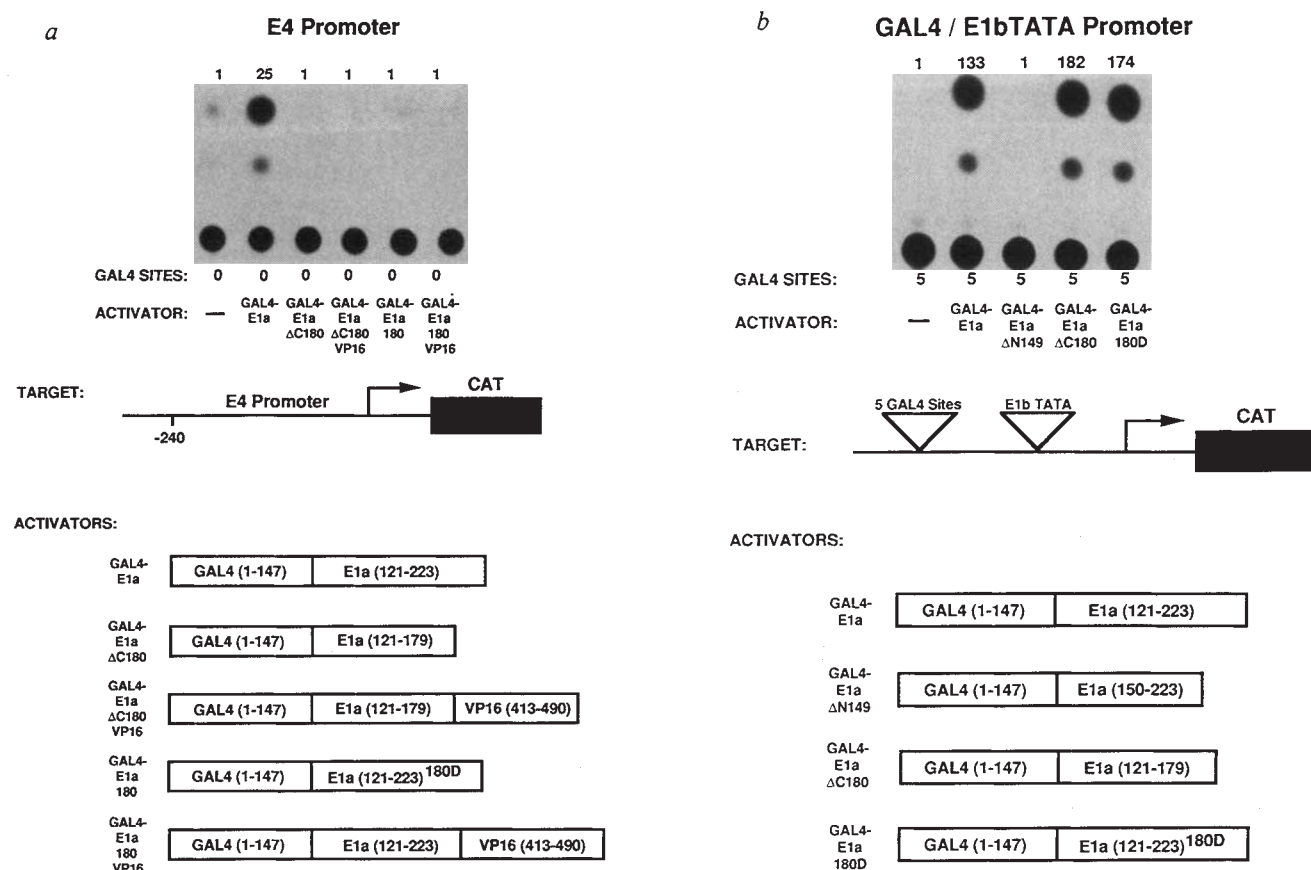


FIG. 5 The C terminus of E1a region 3 provides a promoter binding activity. A reporter CAT plasmid containing the E4 (a) or GAL4-E1bTATA (b) promoter was cotransfected with a plasmid encoding the activator indicated below each track. Relative CAT activity was quantitated and is indicated above each track.

METHODS. Transfections were as described in the Fig. 1 legend. Each activator is a derivative of GAL4-E1a described in the Fig. 1 legend. GAL4-E1aΔN149 is described in the Fig. 2 legend. GAL4-E1aΔC180 is E1a amino acids 121–179

fused to GAL4(1–147) and was created by a 3' *Bal*/31 deletion of the *Clal*–*Xba*I fragment of the E1a 13S cDNA. GAL4-E1a180D contains a substitution at E1a amino acid 180 (Gly→Asp) and was constructed by replacing the *Clal*–*Xba*I fragment of GAL4-E1a with the *Clal*–*Xba*I fragment of the E1a activation mutant pH3G-13S-1098 (ref. 26). GAL4-E1aΔC180-VP16 and GAL4-E1a180D-VP16 contain the 78 C-terminal amino acids of the HSV1 protein fused in-frame to the C terminus of GAL4-E1aΔC180 and GAL4-E1a180D, respectively.

(Fig. 1a, GAL4-VP16). The only other sequences contained in GAL4-E1aΔN149 are E1a amino acids 150–223, suggesting that these E1a sequences provide a promoter-binding function. Below, we confirm this conclusion and further show that the sequences involved in promoter binding are required for E1a's natural activity.

Previous studies have shown that the C-terminal fragment of region 3 is important for wild-type E1a function^{24,26,28,30}. The experiments in Fig. 5 indicate that the primary activity of this region is to bring the E1a activating region to the vicinity of its natural target promoters. Deletion of E1a amino acids 180–223, or substitution of amino acid 180, abolishes GAL4-E1a's ability to activate the E4 promoter, which lacks GAL4 sites (Fig. 5a, GAL4-E1aΔC180, GAL4-E1a180D; E4 promoter). The GAL4 DNA-binding domain restores the activity of these C-terminal mutants but only on a promoter that contains GAL4 sites (Fig. 5b, GAL4-E1aΔC180, GAL4-E1a180D; GAL4/E1bTATA promoter). Thus, the function of E1a amino acids 180–223 can be provided by the binding region of another activator.

The VP16 activating region, which restores the activity of region 3 N-terminal mutants, does not substitute for the C-terminal portion of region 3 (Fig. 5a, GAL4-E1aΔC180-VP16, GAL4-E1a180D-VP16). Conversely, the GAL4 DNA-binding domain, which restores the activity of region 3 C-terminal mutants, does not substitute for the N-terminal portion of region 3 (Fig. 5b, GAL4-E1aΔN149). Thus, the two ends of region 3 provide distinct activities, both of which are required for activation by wild-type E1a.

Model

Taken together our results suggest that E1a functions naturally in the vicinity of the promoter. E1a contains both a promoter-binding region that directs E1a to its target promoters, and an activating region that functions when positioned at the promoter via the GAL4 DNA-binding domain or E1a's natural binding region (Fig. 6a). The activating and binding regions of cellular activators are modular, can be replaced by the corresponding region of an unrelated activator, and are functional independent of their position (reviewed in ref. 3). Similarly, we show that the N-terminal E1a activating region can be replaced by the VP16 activating region at the C-terminus, and the C-terminal E1a binding region can be replaced by the GAL4 DNA-binding domain at the N terminus. The activating regions of typical cellular activators are acidic (reviewed in ref. 3), and we note that the E1a activating region contains several acidic amino acids (Fig. 6c). The central portion of region 3 (amino acids 154–174), which contains a metal-binding domain⁴⁶, is important for wild-type E1a activity^{30,46}. Future experiments will determine if this domain is a component of the activating and/or the promoter-binding regions.

How is E1a targeted to the E4 and other viral early promoters? Our results indicate that E1a functions at the promoter, but there is no evidence that E1a binds to a specific DNA sequence (reviewed in ref. 17). It therefore seems probable that E1a's binding region recognizes a DNA-bound protein and not DNA. In this regard, E1a would resemble activating proteins such as

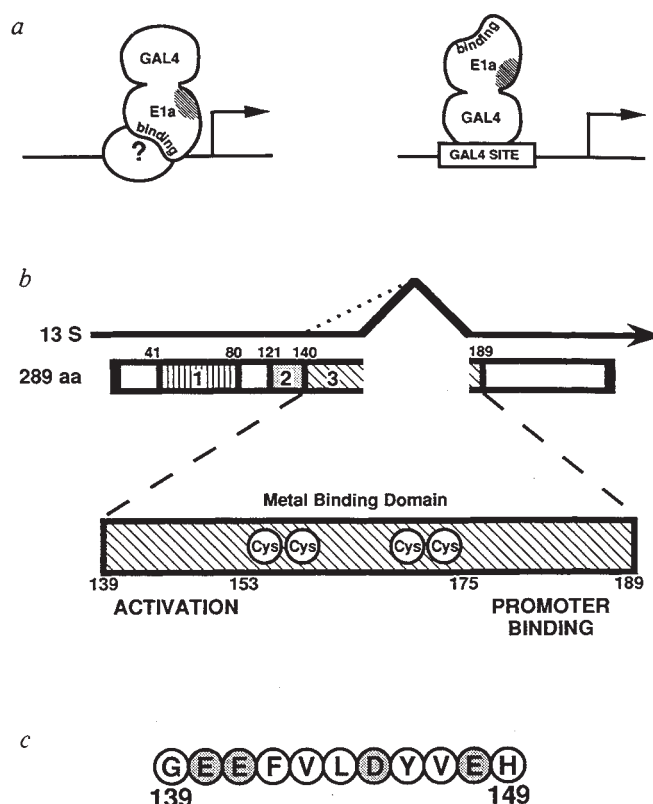


FIG. 6 Model of E1a action. *a*, Transcription activation by GAL4-E1a. A proposal for how GAL4-E1a interacts with target promoters that contain (right) or lack (left) GAL4-binding sites. GAL4-E1a protein is shown as a fusion of GAL4(1–147) and E1a. 'Binding' indicates the E1a promoter-binding region and shading indicates the E1a activating region. The unidentified component with which the E1a binding region interacts in the absence of GAL4 sites is shown by a question mark. *b*, Regions responsible for E1a activity. The structure of E1a 13S mRNA, which encodes the 289-amino acid protein, is shown on top. The intron removed by splicing is designated by a caret. The position of the 12S 5' splice site is indicated by a dotted line. The three conserved regions of the 289-amino acid E1a protein (box) are numbered 1, 2 and 3. The regions involved in activation are shown below. Four of the cysteines within the metal-binding region are shown. *c*, Amino-acid sequence shown to be involved in the E1a activating region. E1a amino acids 139 to 149 are shown. Negatively charged amino acids are stippled.

c-fos and HSV1 VP16, which interact with DNA-bound proteins and contribute additional activating regions^{44–46}. One potential target protein is the cellular transcription factor ATF, which binds to multiple E1a-inducible adenoviral early promoters, including E4 (ref. 47).

If E1a functions at the promoter like a typical cellular activator, what accounts for its promiscuity? One possibility is that E1a's binding region recognizes more than one target protein. In fact, E1a is known to interact with multiple cellular proteins^{48,49}. Perhaps E1a binds to some proteins, like ATF, avidly and to other cellular proteins with lower affinity. This could account for the variable effect of E1a on different pro-

motors¹⁷. In this regard, insertion of GAL4 sites greatly enhances activation of promoters that are otherwise weakly induced by GAL4-E1a, and modestly enhances activation of promoters that are otherwise efficiently induced by GAL4-E1a (Fig. 1; compare E4, MMTV and E1bTATA promoters). Alternatively, there may be additional mechanisms, unrelated to the one described here, by which E1a activates transcription.

In addition to adenovirus, other viruses such as pseudorabies virus⁵⁰, bovine papilloma virus⁵¹ and human T lymphotropic virus-I⁵² encode proteins that activate transcription somewhat promiscuously. We speculate that these viral activators function in a similar way to E1a. □

Received 18 October 1988; accepted 27 January 1989.

- Brent, R. & Ptashne, M. *Cell* **43**, 729–736 (1985).
- Struhl, K. *Cell* **49**, 295–297 (1987).
- Ptashne, M. *Nature* **335**, 683–689 (1988).
- Bohmann, D., Tjian, R. & Franza, R. *Science* **240**, 1010–1016 (1988).
- Chiu, R. et al. *Cell* **54**, 541–552 (1988).
- Sassone-Corsi, P., Lamph, W. W., Kamphaus, M. & Verma, I. M. *Cell* **54**, 553–560 (1988).
- Triebenberg, S. J., LaMarco, K. L. & McKnight, S. L. *Genes Dev.* **2**, 730–742 (1988).
- Preson, C. M., Frame, M. C. & Campbell, M. E. M. *Cell* **52**, 425–434 (1988).
- O'Hare, P. & Goding, C. R. *Cell* **52**, 435–445 (1988).
- Gerster, T. & Roeder, R. G. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6347–6351 (1988).
- Ma, J. & Ptashne, M. *Cell* **48**, 847–853 (1987).
- Ma, J. & Ptashne, M. *Cell* **51**, 113–119 (1987).
- Hope, I. & Struhl, K. *Cell* **46**, 885–894 (1986).
- Hope, I., Subramony, M. & Struhl, K. *Nature* **333**, 635–640 (1988).
- Berk, A. J., Lee, F., Harrison, T., Williams, J. & Sharp, P. A. *Cell* **17**, 935–944 (1979).
- Jones, N. & Shenk, T. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3665–3669 (1979).
- Berk, A. J. *Rev. Genet.* **20**, 45–80 (1986).
- Ferguson, B. et al. *Molec. cell. Biol.* **5**, 2653–2661 (1985).
- Kovesdi, L., Reichel, R. & Nevins, J. R. *Cell* **45**, 219–228 (1986).
- Kovesdi, L., Reichel, R. & Nevins, J. R. *Science* **231**, 719–722 (1986).
- Reichel, R., Kovesdi, L. & Nevins, J. R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 387–390 (1988).
- Leong, K., Brunet, L. & Berk, A. J. *Molec. cell. Biol.* **8**, 1765–1774 (1988).
- Moran, E. & Mathews, M. B. *Cell* **48**, 177–178 (1987).
- Glenn, G. M. & Ricciardi, R. P. *J. Virol.* **56**, 66–74 (1985).
- Moran, E., Zerler, B., Harrison, T. M. & Mathews, M. B. *Molec. cell. Biol.* **6**, 3470–3480 (1986).
- Lillie, J. W., Green, M. & Green, M. R. *Cell* **46**, 1043–1051 (1986).
- Lillie, J. W., Loewenstein, P. M., Green, M. R. & Green, M. *Cell* **50**, 1091–1100 (1987).
- Schneider, J. F., Fisher, F., Goding, C. R. & Jones, N. B. *EMBO J.* **6**, 2053–2060 (1987).
- Green, M., Loewenstein, P. M., Pusztai, R. & Symington, J. S. *Cell* **53**, 921–926 (1988).
- Jelmsma, T. N. et al. *Virology* **163**, 494–502 (1988).

- Webster, N., Jin, J. R., Green, S., Hollis, M. & Chambon, P. *Cell* **52**, 169–178 (1988).
- Sadowski, I., Ma, J., Triezenberg, S. & Ptashne, M. *Nature* **335**, 559–562 (1988).
- Keegan, L., Gill, G. & Ptashne, M. *Science* **231**, 699–704 (1986).
- Fischer, J. A., Ginger, E., Maniatis, T. & Ptashne, M. *Nature* **332**, 853–856 (1988).
- Ma, J., Przbilla, E., Hu, J., Bogorad, L. & Ptashne, M. *Nature* **334**, 631–633 (1988).
- Kakidani, H. & Ptashne, M. *Cell* **52**, 161–167 (1988).
- Wu, L., Rosser, D. S. E., Schmidt, M. C. & Berk, A. J. *Nature* **326**, 512–515 (1987).
- Slavicek, J. M., Jones, N. C. & Richter, J. D. *EMBO J.* **7**, 3171–3180 (1988).
- Bachmair, A., Finley, D. & Varshavsky, A. *Science* **234**, 179–186 (1986).
- Paine, P. L., Moore, L. C. & Horowitz, S. B. *Nature* **254**, 109–114 (1975).
- Lang, I., Scholz, M. & Peters, R. *J. Cell Biol.* **102**, 1183–1190 (1986).
- Triebenberg, S. J., Kingsbury, R. C. & McKnight, S. L. *Genes Dev.* **2**, 718–729 (1988).
- Godowski, P. J., Picard, D. & Yamamoto, K. R. *Science* **241**, 812–816 (1988).
- Hollenberg, S. M. & Evans, R. *Cell* **55**, 899–906 (1988).
- Lech, K., Anderson, K. & Brent, R. *Cell* **52**, 179–184 (1988).
- Culp, J. S. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6450–6454 (1988).
- Lee, K. A. W. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8355–8359 (1987).
- Harlow, E., Whyte, P., Franza, B. R. & Schley, C. *Molec. cell. Biol.* **6**, 1579–1589 (1986).
- Yee, S.-P. & Branton, P. *Virology* **147**, 141–153 (1985).
- Feldman, L. T., Imperiale, M. F. & Nevins, J. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4952–4956 (1982).
- Haugen, T. H., Cripe, T. P., Ginder, G. D., Karin, M. & Turek, L. P. *EMBO J.* **6**, 145–152.
- Ballard, D. W. et al. *Science* **241**, 1652–1656 (1988).
- Svensson, C., Pettersson, U. & Akusjärvi, G. *J. molec. Biol.* **165**, 475–499 (1983).
- Velich, A. & Ziff, E. *Cell* **40**, 705–716 (1985).
- Cato, A., Miksic, R., Schuz, G., Arnemann, J. & Beato, M. *EMBO J.* **5**, 2237–2240 (1986).
- Gorman, C. M., Moffat, L. F. & Howard, B. H. *Molec. cell. Biol.* **2**, 1044–1051 (1982).

ACKNOWLEDGEMENTS. We thank Ivan Sadowski, Kathy Martin, Roger Brent, Mark Ptashne and members of the Ptashne laboratory for advice and suggestions, and Ivan Sadowski for gifts of plasmids and antibodies. This work was supported by grants from the NIH and The Chicago Community Trust/Searle Scholars Program to M.R.G.

Bars within bars: a mechanism for fuelling active galactic nuclei

Isaac Shlosman*, Juhan Frank†
& Mitchell C. Begelman‡

Joint Institute for Laboratory Astrophysics, University of Colorado and
National Bureau of Standards, Boulder, Colorado 80309, USA

ACTIVE galactic nuclei (AGNs) are usually thought to be powered by accretion onto a supermassive black hole (SBH)^{1,2}. Their luminosities, which may exceed 10^{46} erg s⁻¹, require mass accretion rates of $\geq 1 M_{\odot}$ yr⁻¹ for reasonable mass-to-energy conversion efficiencies. Although this quantity of fuel could be supplied by the interstellar medium of the host galaxy, it is not obvious how it could be transported from typical galactic radii, ~ 10 kpc, down to the scale of the SBH, ≤ 10 pc. We propose here a mechanism, applicable to AGNs and nuclear starburst galaxies, which brings in gas from large to small scales by successive dynamical instabilities. On the large scale, a stellar bar sweeps the interstellar medium into a gaseous disk of a few hundred parsecs in radius. Under certain conditions, this disk can become unstable again, allowing material to flow inwards until turbulent viscous processes control angular-momentum transport. This flow pattern may feed viscosity-driven accretion flows around a SBH, or lead to the formation of a SBH if none was present initially.

If the accretion flow is disk-like and is driven by local viscosity, then the radial inflow time can be expressed in terms of the standard α parameter³, which is defined as a suitably averaged ratio of viscous stress to pressure inside the disk. For a disk with shear induced by a point-mass potential, the viscosity-controlled radial inflow time, t_{visc} , is

$$t_{\text{visc}} = \alpha^{-1} (v_{\phi}/c_s)^2 t_{\text{dyn}} \approx 1.2 \times 10^9 \alpha^{-1} v_{\phi 2} T_2^{-1} R_{10} \text{ yr} \quad (1)$$

where $100 v_{\phi 2}$ km s⁻¹ is the orbital speed, $10 R_{10}$ pc is the radius, t_{dyn} is the dynamical time R/v_{ϕ} , c_s is the sound speed in the disk, and $100 T_2$ K is the gas temperature^{4,5}. An alternative viscosity formulation, obtained by setting the coefficient of kinematic viscosity to $\nu = \alpha c_s H$, where H is the disk thickness, gives identical results for a non-self-gravitating disk, but leads to a slower inflow rate when vertical self-gravity is important. Because of the effects of self-gravity and fragmentation⁴⁻⁷, the disk feeding an AGN may resemble a system of gas clouds rather than a quasi-continuous fluid. In this case, the efficiency of angular-momentum transfer depends on the cloud-cloud collision timescale, t_{coll} , and the value of the effective α is of order $t_{\text{dyn}}/t_{\text{coll}}$. We expect α to be ≤ 1 at the onset of fragmentation, and to become smaller if the clouds contract further. If fragmentation is somehow suppressed, higher effective values of α may be possible⁸, but a consideration of realistic heating and cooling processes suggests that this does not lead to a significant enhancement of inflow^{4,5}. In regions where the stellar potential dominates over the potential of the SBH, the shear is reduced and the viscous inflow time is prolonged. We conclude, therefore that standard viscous processes are too slow to be effective in collecting gas liberated in the main body of the host galaxy. This can be overcome either if the gas is generated close to the SBH or if gas liberated at larger distances is brought in towards the SBH by a mechanism other than subsonic turbulent viscosity. For example, a large-scale non-axisymmetric disturbance of the galaxy, such as a stellar bar, is capable of inducing global shocks, thereby driving the inflow of gas towards the

centre. The amount of H I in spirals is observed to lie in the range 10^8 – $10^{10} M_{\odot}$, and molecular gas may be more abundant by a factor of ten. Both the total H I, CO and H₂ content and the ratio of molecular gas to H I are maximal for spiral galaxies of type Sb to Sbc^{9,10}. There is certainly enough gas present in spirals to power the whole range of AGN activity, provided that this gas can be swept inwards from several kpc down to tens of pc in a short time.

It has been known for some time that barred spiral galaxies are associated with Seyfert galaxies. Adams¹¹ noticed that a large number of Seyferts are barred. The 12 objects in his highest-resolution class are all barred spirals at some level, or show inner rings thought to be generated by bar instabilities¹². Apart from ten obvious SB and SAB galaxies, this class contains NGC1068 and Mk10, which are observed to be barred¹³⁻¹⁵. In the next resolution class, Adams found that two thirds of the galaxies are SB or have rings. The remaining third either have an "amorphous body" or are seen edge-on, or their observation was subject to instrumental errors.

In a survey of a sample of Seyfert galaxies, Simkin *et al.*¹⁶ concluded that there was an excess of bars, rings and oval distortions among Seyferts with respect to their control sample. A recent survey based on CCD images of a volume- and luminosity-limited sample by MacKenty¹⁷ shows that Seyfert galaxies "nearly always possess mechanisms for transporting material into their nuclei (for example, peculiar, tidally interacting or barred galaxies)". In other respects the host galaxies appear normal, having colours, magnitudes and disk parameters typical of their morphological class. Clearly, the simplest interpretation of these observations is that all Seyferts are barred at some level.

The large-scale stellar bar cannot, however, be solely responsible for the AGN phenomenon. First, about half of the spiral galaxies in the sky are barred or show oval distortions whereas only a few per cent of them host AGNs (but see ref. 18). Second, the gas inflow generated by a barred rotating potential does not extend down to scales at which turbulent viscosity could take over. Numerical studies of the response of a compressible fluid to an imposed oval distortion of the potential indicate that, in these systems, the gaseous component loses angular momentum in a pair of shocks, and flows towards the centre of the galaxy in about ten rotation periods^{19,20}. Given the fractional deviation of the potential from axial symmetry, $\Delta\Phi$, we can estimate the radius at which the shocks weaken and the inflow ceases to be dynamically driven. We assume that $\Delta\Phi \approx \epsilon v_{\phi}^2$, where ϵ is of the order of 0.1. The condition for $\Delta\Phi$ to cause a shock in the gas can be expressed as²¹ $c_g < 0.5(\gamma+1)\epsilon v_{\phi}$, where c_g is the sound speed in the gas and γ is its adiabatic index. If the inner parts of the galaxy rotate as a solid body within a radius R_i , shocks will not be present within a radius R_d given by $R_d = 2R_i/(\gamma+1)\epsilon\mathcal{M}$, where $\mathcal{M}$ is the Mach number (v/c_g) of the flow at R_i and we have assumed $c_g \ll v_{\phi}$. Numerical simulations show that the gas can be swept up into a ring if inner Lindblad resonances (ILRs) are present, yielding a ring-like configuration of H II regions. If the bar distortion is strong, however, the positions of the ILRs cannot be calculated using the circular velocity curve²².

In summary, the gaseous inflow in a barred potential slows down at a radius R_d . We shall assume, for simplicity, that the net effect of a stellar bar on the gaseous component of the host galaxy is to sweep it into a disk with radial scale $R_d \approx 0.1 R_b$, where R_b is the radius of the bar. At R_d , viscous processes are still too slow to drive inflow. We propose instead that the gas which accumulates in the central kpc or so, as a consequence of the inflow in the stellar bar, forms a disk which, under certain conditions, becomes again dynamically unstable to form a gaseous bar. We derive a condition for this secondary instability in the context of a simple model, and argue that the resulting inflow can extend all the way into the centre of the galaxy, that is, to the inner ~ 10 pc or so.

No single criterion is known for global stability of self-gravi-

* Present address: Theoretical Astrophysics 130-33, California Institute of Technology, Pasadena, California 91125, USA.

† Permanent address: Max-Planck-Institut für Astrophysik, 8046 Garching bei München, FRG.

‡ Also at: Department of Astrophysical, Planetary, and Atmospheric Sciences, University of Colorado, Boulder, Colorado 80309, USA.

tating disks (gaseous or stellar) encompassing all possible disk models^{22,23}. However, the stability of a large variety of disk models can be described by applying the semi-empirical criterion proposed by Ostriker and Peebles²⁴, which is known to have a theoretical analogue for Maclaurin ellipsoids. A sufficient condition for instability is given by $t = T_{\text{rot}}/|W| \geq t_{\text{crit}}$, where T_{rot} is the kinetic energy of rotation of the disk, W is the gravitational energy of the system and t_{crit} is the maximum value of t required for stability. This criterion seems to apply to gaseous thin disks, which have $t_{\text{crit}} \approx 0.14$ for secular instability and $t_{\text{crit}} \approx 0.26$ for dynamical instability^{25,26}. We model the system after the bar has brought the gas within the radius R_d in terms of three components: a cool gaseous disk having a Kuzmin density distribution with radial scale a ; a stellar disk having a similar density distribution but with a larger radial scale b ; and a Plummer-model halo, also with radial scale b (see, for example, ref. 23). Let the total mass of the system be M , a fraction g of which is in the gaseous disk and a fraction s in the stellar disk. The halo therefore contains a mass $(1 - g - s)M$. This particular choice of model is motivated by the fact that t can be evaluated analytically, but the conclusions drawn from it are expected to be at least qualitatively correct in a more general context.

We treat the potential of the stellar component as axisymmetric and fixed, and therefore include in the calculation of t only the kinetic energy of the gaseous disk:

$$T_{\text{rot}} = \frac{gf^2 GM^2}{8a} \left[g + \frac{4(1-g)}{(1+\beta)^2} \right] \quad (2)$$

where we have assumed that the disk rotates at every radius with a fraction f of the local circular speed for equilibrium ($f=1$ indicates full rotational support), and have introduced $\beta \equiv b/a$, the ratio of length scales of the fixed stellar component to the gaseous component. From equation (2) and the expression for the potential energy,

$$|W| = \frac{GM^2}{4a} \left[g^2 + \frac{4g(1-g)}{1+\beta} + \frac{s(2-2g-s)}{\beta} + \frac{3\pi}{8\beta} (1-s-g)^2 \right] \quad (3)$$

one can obtain an expression for t in terms of the mass fractions in the various components and the radial-scale ratio β . The physically interesting question is: given a certain gas fraction, by how much must the gas disk shrink radially in order to become unstable? A rough answer to this question can be obtained by setting t equal to t_{crit} and solving for β . In the limit $g \ll 1$, we find that the gaseous disk is unstable ($t > t_{\text{crit}}$) if

$$\beta > \frac{2t_{\text{crit}}(1-h^2+3\pi h^2/8)}{g^2(f^2-2t_{\text{crit}})} \quad (4)$$

where $h \equiv 1 - g - s$ is the halo mass fraction. The r.h.s. of eq. (4) is not very sensitive to the value of h , and equation (4) holds to better than 10% for $g \leq 0.3$, when compared with the exact evaluation of the Ostriker-Peebles criterion. Any effects of non-axisymmetry will probably enhance the instability.

Two aspects of equation (4) are worth mentioning: (1) for a given f and h , β varies inversely with the gas fraction squared, and (2) the disk must be cold, $f > \sqrt{2t_{\text{crit}}}$ for it ever to become unstable. This places stringent limits on the proposed mechanism. For example, in a cold disk with $f=1$, $h \approx 0.5$ and $t_{\text{crit}} = 0.14$, β must be $> 0.4/g^2$ for instability. If $\beta=1$ initially and $g \approx 0.2$, then a reduction in size of the gas disk by a factor of ten would make it susceptible to global instabilities. It is important to note that the value of $t_{\text{crit}} = 0.14$ used above corresponds to dynamical instability of stellar (collisionless) disks, but to only secular instability of gaseous disks. This is because gaseous disks have fewer degrees of freedom than stellar disks. Secular instability of a gaseous disk would occur on the viscous inflow time, and would therefore be too slow to be of interest. Dynamical instability of a smooth gaseous disk would require $t_{\text{crit}} = 0.26$, corresponding either to contraction by an additional factor of 2-3,

or to increase in g by a factor of about two. However, it is likely that the self-gravitating gaseous disk will be highly inhomogeneous, consisting of clouds which fill a rather small fraction of the volume⁵. If the collision frequency of clouds is small relative to the orbital frequency, the system will behave more like a stellar disk than a gaseous disk, and a lower value of t_{crit} will apply. Gravitational coupling of the gaseous disk to the stellar disk may also contribute to the instability^{5,27}.

Therefore, we have established that a large-scale stellar bar sweeping the gas inward yields a disk, which can become unstable, if it is not too hot, and can feed a viscosity-driven flow. The unstable configuration is likely to resemble a bar-driven spiral²⁵, which will transfer angular momentum outwards, leading to further contraction. At the same time, the velocity dispersion of inhomogeneities will tend to increase, but this energy will be dissipated through cloud-cloud collisions. To the extent that the system does not form stars, the inevitable consequence of sinking of the gaseous system towards the centre of the stellar system is that t will grow, and the gaseous system will become increasingly unstable. The degree to which the system forms stars before the gas reaches the centre must, in this model, determine whether the principal form of activity is a starburst or an AGN. The above argument implies that galaxies with a gas content of less than 10% will not display a bar within a bar and are therefore candidates for the generation of a starburst but not of a powerful AGN.

The process is not necessarily restricted to spirals. Elliptical galaxies are to some extent also supported by rotation. Material that is released at radii of a few kpc, and which is only 1% rotationally supported, would go into keplerian orbit at about one tenth of its original distance from the centre. Turbulent viscosity would still be insufficient to drive inflow, and the formation of a gaseous bar might occur even in this case.

Thus, a large-scale bar is a necessary but not sufficient condition for strong nuclear activity if the host is a spiral galaxy. Another stage of dynamical instability, which occurs in the gaseous disk that accumulates in the centre under the influence of the stellar bar, is required to feed an accretion disk at smaller scales, that is, such a process requires gaseous bars within stellar bars. A similar but tidally induced process was found in numerical simulations (L. Hernquist, preprint). Any process that suppresses this second stage of instability will prevent formation of the SBH and/or its subsequent fuelling, and is likely to lead to nuclear star formation or to low-level activity. This view is supported by recent evidence that most or perhaps all starbursts are barred²⁸. The majority of spirals avoids becoming a conspicuous AGN or starburst. The most probable causes for this are relatively low gas content of the host, inefficiency of the stellar bar in sweeping all of the gas in, and the presence of ILRs.

The fact that activity seems to occur preferentially in intermediate morphological classes can be simply explained in terms of the relative abundance of gas-rich galaxies among those morphological types^{9,10}. Additional factors, such as the length of the bar relative to the photometric galactic radius, may influence the correlation between activity and morphological type. It is plausible that nuclear activity in some galaxies may be suppressed at present because the SBH (or the nuclear starburst) has exhausted the available mass supply and must wait for new material to be brought in from large scales. The duration of the active phase of an AGN would be of the order of a few t_{dyn} at the bar radius, whereas the gas would be replenished only on the stellar-evolution timescale or on the timescale for capture of extragalactic clouds.

An important prediction of this model is that dynamical instabilities on scales of ~ 100 pc must be present in the central regions of active galaxies, including ellipticals (such as powerful radio galaxies). Perhaps some evidence of the processes discussed here can be found in the Sérsic-Pastoriza galaxies, whose nuclei have been long recognized as unusually bright and disturbed²⁹. All of these galaxies are barred at some level and some

of them are well-known today to be Seyferts and starbursts. More high-resolution optical and infrared observations of these objects and H I and CO observations of the innermost regions of moderately distant AGN are desirable.

We have not discussed the effects of environment on Seyferts and starbursts. The passage of a close companion may induce a stellar bar, but this may not be necessary for their production, as isolated galaxies may host bars which are relics from galaxy formation, or which have been generated spontaneously as a result of infall. □

Received 29 July 1988; accepted 9 January 1989.

- Lynden-Bell, D. *Nature* **223**, 690–694 (1969).
- Begelman, M. C., Blandford, R. D. & Rees, M. J. *Rev. mod. Phys.* **56**, 255–351 (1984).
- Shakura, N. I. & Sunyaev, R. A. *Astr. Astrophys.* **24**, 337–355 (1973).
- Shlosman, I. & Begelman, M. C. *Nature* **329**, 810–812 (1987).
- Shlosman, I. & Begelman, M. C. *Astrophys. J.* (in the press).
- Paczynski, B. *Acta Astr.* **28**, 91–109 (1978).
- Shore, S. N. & White, R. L. *Astrophys. J.* **256**, 390–396 (1982).
- Lin, D. N. C. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **225**, 607–613 (1987).
- Haynes, M. P. & Giovanelli, R. *Astr. J.* **89**, 758–800 (1984).
- Verter, F. *Astrophys. J.* **65**, 555–580 (1987).
- Adams, T. F. *Astrophys. J. Suppl.* **33**, 19–34 (1977).
- Buta, R. *Astrophys. J. Suppl.* **61**, 609–630 (1986).
- Telesco, C. M. & Decher, R. *Astrophys. J.* (submitted).
- Scoville, N., Matthews, K., Carico, D. P. & Sanders, D. B. *Astrophys. J.* **327**, L61–L64 (1988).
- Mazzarella, J. M. & Balzano, V. A. *Astrophys. J. Suppl.* **62**, 751–819 (1986).
- Simkin, S., Su, H. & Schwarz, M. P. *Astrophys. J.* **237**, 404–413 (1980).
- MacKenty, J. *Astrophys. J.* (in the press).
- Keel, W. C. *Astrophys. J. Suppl.* **52**, 229–239 (1983).
- Matsuda, T., Inoue, M., Sawada, K., Shima, E. & Wakamatsu, K. *Mon. Not. R. astr. Soc.* **229**, 295–314 (1987).
- Schwartz, M. P. *Mon. Not. R. astr. Soc.* **212**, 677–686 (1985).
- Landau, L. D. & Lifshitz, E. M. *Fluid Dynamics* 94 (Pergamon, New York, 1959).
- Athanasoulas, E. *Phys. Rep.* **114**, 329–403 (1984).
- Binney, J. & Tremaine, S. *Galactic Dynamics* (Princeton University Press, 1987).
- Ostriker, J. P. & Peebles, P. J. E. *Astrophys. J.* **186**, 467–480 (1973).
- Bardeen, J. M. in *Dynamics of Stellar Systems* (ed. Hayli, A.) 297–320 (Reidel, Dordrecht, 1975).
- Aoki, S., Noguchi, M. and Iye, M. *Publ. astr. Soc. Japan* **31**, 737–774 (1987).
- Jog, C. J. & Solomon, P. M. *Astrophys. J.* **276**, 114–126 (1984).
- Jackson, J. M., Barrett, A. H., Armstrong, J. T. & Ho, P. T. P. *Astr. J.* **93**, 531–545 (1987).
- Sérsic, J. L. & Pastoriza, M. *Publ. astr. Soc. Pacif.* **77**, 287–289 (1965).

ACKNOWLEDGEMENTS. We are grateful to Oved Dahari, Don Osterbrock, John Papaloizou, John Stoeck and Andrew S. Wilson for illuminating discussions and to Lars Hernquist for providing results before publication. This research was supported in part by NASA and the NSF and grants from Ball Aerospace Systems Division, Rockwell International Corporation and the Alfred P. Sloan Foundation. M.C.B. is a Presidential Young Investigator and Alfred P. Sloan Foundation Research Fellow.

A global assessment of natural sources of atmospheric trace metals

Jerome O. Nriagu

National Water Research Institute, Box 5050, Burlington, Ontario L7R 4A6, Canada

A PROPER inventory of atmospheric emissions from natural sources is basic to our understanding of the atmospheric cycle of the trace metals (and metalloids), and is also needed for assessing the extent of regional and global pollution by toxic metals¹. It is generally presumed that the principal natural sources of trace metals in the atmosphere are wind-borne soil particles, volcanoes, seasalt spray and wild forest fires^{2–6}. Recent studies have shown, however, that particulate organic matter is the dominant component of atmospheric aerosols in non-urban areas^{7–10} and that over 60% of the airborne trace metals in forested regions can be attributed to aerosols of biogenic origin^{11,12}. Here I estimate that biogenic sources can account for 30–50% of the global baseline emissions of trace metals. For most of the toxic metals, the natural fluxes are small compared with emissions from industrial activities, implying that mankind has become the key agent in the global atmospheric cycle of trace metals and metalloids.

In compiling the emission factors used in the calculations (Table 1), the most recent data have been used; for older data, the lowest reported concentrations have been preferred. For

enrichment factors, a range of values¹, rather than a single value, has been used. A range in the flux of material from each source has also been employed so as to assess the errors inherent in the reported emission intensities of trace metals.

The episodic nature of volcanic emissions makes it difficult to derive trace-metal emission rates using the enrichment-factor strategy^{12,13}. There have been numerous studies, however, of the release of sulphur and trace metals from volcanoes and fumaroles, and a number of authors have estimated the outputs of trace metals by normalizing the metal release to the sulphur flux from this source¹⁵. I have adopted such an approach here, using a global volcanic sulphur flux of $(15\text{--}50) \times 10^{12} \text{ g yr}^{-1}$ (refs 15, 16) and metal-to-sulphur ratios derived from published data^{13–17}.

The release of large quantities of volatile non-methane hydrocarbons (NMHC) is well documented, particularly in forested areas^{18–20}. For example, the average emission rates for NMHCs in the forested areas of the United States range from 450 to $1,712 \mu\text{g m}^{-2} \text{ h}^{-1}$ (ref. 18). Biogenic emissions of isoprene and terpenes in the Amazon rain-forests have been estimated as 233 and $1,040 \mu\text{g m}^{-2} \text{ h}^{-1}$ respectively¹⁹. Rasmussen and Khalil²⁰ have calculated the global atmospheric flux of isoprene to be $450 \times 10^{12} \text{ g yr}^{-1}$, and the combined releases of isoprene and terpenes have been suggested¹⁹ to be equivalent to ~0.7% of the net global primary production. Thus the NMHC flux of $100\text{--}500 \mu\text{g m}^{-2} \text{ h}^{-1}$ used here is clearly a conservative figure. Terpenes and isoprene, the two dominant components of NMHCs, are known to form strong complexes with many trace metals²¹, and thus should play a part in the transfer of metals to the atmosphere. As far as I know, however, the NMHCs have not been analysed for trace metals, so that the emission factors listed in Table 1 are subject to a large uncertainty. They are based on reported flux measurements and the concentrations of methylated compounds of Hg, As, Pb and Se in the atmosphere^{6,14,24}. For the other elements, the emission factors are based on metal concentrations in the surface organic microlayers of aquatic ecosystems^{22,23}.

The average concentration of particulate organic carbon (POC; derived from organic matter which is the dominant component of atmospheric aerosols) in the Amazon rain-forests has been reported⁹ to be $8.9 \mu\text{g C m}^{-3}$ in the mixed layer and $2.6 \mu\text{g C m}^{-3}$ in the free troposphere. These reported POC concentrations in forest ecosystems are considerably higher than the mean value of $0.06 \mu\text{g C m}^{-3}$ in the marine atmosphere of the Southern Hemisphere⁴⁴. The carbon isotope composition suggests, however, that most of the marine POC is also derived from natural land-based sources⁴⁴. The annual flux of particulate organic matter (POM) to the atmosphere, Q , can be estimated from

$$Q = V \times B \times 2.2 \times (3\text{--}6) \times 10^{13}$$

where V , the total deposition velocity of the aerosols, is estimated to be 1.0 cm s^{-1} (ref. 44), and B , the average POC concentration, is $3\text{--}9 \mu\text{g m}^{-3}$. The figure of 2.2 is used to relate organic carbon to organic matter⁴⁷, and the forested area is believed to be $(3\text{--}6) \times 10^{13} \text{ m}^2$ (ref. 51). These values yield a global atmospheric POM flux of about $(1\text{--}5) \times 10^{14} \text{ g yr}^{-1}$. The POM flux used here is thus higher than the value of $27 \times 10^{12} \text{ g yr}^{-1}$ reported recently by Cachier *et al.*⁴⁴. The global flux of material from the other non-biogenic sources are summarized in Table 1.

There is a wide range in the estimated total emission of each metal from any given source (Table 1), reflecting the large dispersion in the emission factors used and in the published estimates of the mass flux from individual sources. There are significant gaps in the data, demonstrating the need for further study of the role of biological processes in the emission of trace metals. The proposed limits on the fluxes of metals from each source serve to indicate that current data are sufficient for only order-of-magnitude estimates of the global emission intensities. In view of this, the central (or median) values of the ranges

TABLE 1 Worldwide emissions of trace metals from natural sources ($\times 10^9 \text{ g yr}^{-1}$)

	As	Cd	Co	Cr	Cu	Hg	Mn	Mo	Ni	Pb	Sb	Se	V	Zn
Wind-borne soil particles (0.6–5) $\times 10^{14} \text{ g}$														
Emission factor ($\mu\text{g g}^{-1}$)†	5–10	0.2–0.8	10–15	60–100	15–30	0.05–0.2	700–800	2.0–5.0	30–40	5.0–15	1.0–3.0	0.1–0.7	20–60	50–80
Amount emitted: Range	0.3–5.0	0.01–0.4	0.6–7.5	3.6–50	0.9–15	0–0.1	42–400	0.12–2.5	1.8–20	0.3–7.5	0.06–1.5	0.01–0.35	1.2–30	3.0–35
Median value	2.6	0.21	4.1	27	8.0	0.05	221	1.3	11	3.9	0.78	0.18	16	19
Seasalt spray§ (0.1–1.0) $\times 10^{16} \text{ g}$														
Emission factor ($\mu\text{g kg}^{-1}$)	186–314	2.8–11	1.4–14	30–140	230–690	0.14–4.3	20–170	8.6–4.3	93–260	115–285	28–114	43–114	143–714	17–86
Amount emitted: Range	0.19–3.1	0–0.11	0–0.14	0.03–1.4	0.23–6.9	0–0.04	0.02–1.7	0.01–0.43	0.01–2.6	0.02–2.8	0–1.1	0–1.1	0.14–7.2	0.02–0.86
Median value	1.7	0.06	0.07	0.07	3.6	0.02	0.86	0.22	1.3	1.4	0.56	0.55	3.1	0.44
Volcanoes¶ (15–50) $\times 10^{12} \text{ g S}$														
Metal/S ratio ($\times 10^{-4}$)	0.1–1.5	0.09–0.3	0.01–2.5	0.5–5.8	0.6–3.5	0.02–0.4	2.8–16	0.03–0.15	0.62–5.6	0.36–1.2	0–0.27	0.07–0.35	0.14–2.1	0.21–3.8
Amount emitted: Range	0.15–7.5	0.14–1.5	0.02–1.9	0.81–29	0.9–18	0.03–2.0	4.2–80	0.04–0.75	0.93–28	0.54–6.0	0.01–1.4	0.10–1.8	0.21–11	0.31–19
Median value	3.8	0.82	0.96	15	9.4	1.0	42	0.4	14	3.3	0.71	0.95	5.6	9.6
Wild forest fires** (2–15) $\times 10^{14} \text{ g}$														
Emission factor ($\mu\text{g g}^{-1}$)††	0–0.25	0–0.15	0.1–0.4	0–0.12	0.5–50	0–0.3	6.0–30	0.2–0.75	0.5–3.5	0.1–2.5	0–0.3	0–0.35	0.1–2.4	1.5–10
Amount emitted: Range	0–0.38	0–0.22	0.02–0.6	0–0.18	0.1–7.5	0–0.05	1.2–45	0.04–1.1	0.1–4.5	0.06–3.8	0–0.45	0–0.52	0.02–3.6	0.3–15
Median value	0.19	0.11	0.31	0.09	3.8	0.02	23	0.57	2.3	1.9	0.22	0.26	1.8	7.6
Biogenic														
Continental particulates‡‡ (1–5) $\times 10^{14} \text{ g}$														
Emission factor ($\mu\text{g g}^{-1}$)§§	0.2–1.0	0–0.6	0.5–2.0	1.0–4.0	1.0–10	0–0.8	20–100	0.4–1.5	1.0–6.0	0.2–5.0	0–0.6	0–0.5	0.2–3.5	3–10
Amount emitted: Range	0.2–0.5	0–0.83	0.05–1.0	0.1–2.0	0.1–5.0	0–0.04	4–50	0.04–0.75	0.1–1.0	0.02–2.5	0–0.4	0–0.25	0.04–1.8	0.3–5.0
Median value	0.26	0.15	0.52	1.0	2.6	0.02	27	0.40	0.51	1.3	0.2	1.12	0.92	2.6
Continental volatiles (2–25) $\times 10^{13} \text{ g}$														
Emission factor ($\mu\text{g g}^{-1}$)¶¶	1.0–10	0–0.3	0–0.5	0–0.4	0.3–2.5	0.5–5.0	1.0–10	0–0.5	0.1–0.8	0.2–1.5	0–0.3	5–20	0.3–1.0	0.5–20
Amount emitted: Range	0.03–2.5	0–0.8	0–0.12	0–0.10	0.01–0.62	0.02–1.2	0.03–2.5	0–0.12	0–0.20	0.01–0.38	0–0.8	0.15–5.0	0.01–0.25	0.02–5.0
Median value	1.3	0.04	0.06	0.05	0.32	0.61	1.3	0.06	0.10	0.20	0.04	2.6	0.13	2.5
Marine¶¶¶ (8–30) $\times 10^{13} \text{ g}$														
Emission factor ($\mu\text{g g}^{-1}$)¶¶¶	2.0–15	0–0.3	0–0.5	0–0.4	0.3–2.5	0.5–5.0	1.0–10	0–0.5	0.1–0.8	0.2–1.5	0–0.3	5–30	0.3–1.0	0.5–20
Amount emitted: Range	0.16–4.5	0–0.1	0–0.15	0–0.12	0.02–0.75	0.04–1.5	0.08–3.0	0–0.15	0.01–0.45	0.02–0.45	0–0.1	0.4–9.0	0.02–0.3	0.04–6.0
Median value	2.3	0.05	0.08	0.06	0.39	0.77	1.5	0.08	0.12	0.24	0.05	4.7	0.16	3.0
Total emission: Range	0.86–23	0.15–2.6	0.69–11	4.5–83	2.3–54	0.1–4.9	52–582	0.14–5.8	3.0–57	0.97–23	0.07–4.7	0.66–18	1.6–54	4.0–86
Median value	12	1.3	6.1	44	28	2.5	317	3.0	30	12	2.4	9.3	28	45

The amount of each metal emitted is derived by multiplying its emission factor by the flux of material from the particular source.

† The range in mass of wind-borne soil particles was reported by Prospero *et al.*³²

‡ The ranges in average trace-metal concentration of the soil are based on the compilations and reviews by Bowen³³, and Ure and Berrow³⁴.

§ Based on estimates by Blanchard³⁵.

|| Based on the recently reported concentrations of trace metals in sea water^{14,36–41}, an average seawater salinity of 35‰, and an enrichment factor for sea salt of 10 for Mn, Mo, Cr and V, 50 for Se, Sb, Hg, Co, Ni, Cd and Ni, and 200 for Cu, Zn, Pb and Cd^{28,42,43}.

¶ From Berresheim and Jaeschke¹⁶ and Lambert *et al.*¹⁵.

¶¶ Based on trace-metal concentrations in volcanic emanations reported in refs 5, 13, 15–17 and 44.

** Quantity of forest biomass consumed by wild fires derived from the data given in refs 45–47.

†† Derived from the trace-metal concentrations in land plants^{4,13} and assuming burning yields of 30% for Mn and Cr, 50% for Pb, Co, Cd, Mo, Sb, Cu and Zn, and 70% for Se, Hg, V and As (refs 46, 49).

‡‡ See text.

§§ Trace-metal concentrations in land plants, which are used to derive the emission factors in††.

||| The average emission of volatile non-methane natural hydrocarbons (NMHC) in the United States has been estimated to be 450, 884 and 1,712 $\mu\text{g m}^{-2} \text{ h}^{-1}$, respectively (refs 25–27). My calculation is based on NMHC flux of only (100–500) $\mu\text{g m}^{-2} \text{ h}^{-1}$ from (3–6) $\times 10^{13} \text{ m}^2$ of forested area^{45,50}.

¶¶ The data listed for As, Hg and Se are based on published measurements of volatilization rates (for review see, for example, refs 3, 14, 22, 23, 30, 31). The emission factors reported for the other metals are based on their concentrations in surface organic microlayers.

¶¶¶ I have assumed that the flux of volatile NMHCs from the 170 $\times 10^{12} \text{ m}^2$ of coastal shelf and upwelling areas of the ocean⁵¹ is 50–200 $\mu\text{g m}^{-2} \text{ h}^{-1}$.

given should be used in preference to the mean values used by other authors.

Based on these data, my inventories suggest that biogenic sources can account, on average, for over 50% of the Se, Hg and Mo, and 30–50% of the As, Cd, Cu, Mn, Pb and Zn, released annually to the atmosphere from natural sources (Table 1). This important source of metals in the atmosphere has been either ignored or badly underestimated in most of the previously published inventories of trace-metal emissions from natural sources (for example, refs 2–6). This large estimated biogenic flux is consistent with recent reports^{10,11} that the vegetation accounts for up to 90% of the airborne trace metals in the Amazon Basin. Biologically mediated volatilization processes can account for 30–50% of the total Hg, As and Se emitted annually, whereas the other metals are primarily associated with pollens, spores, waxes, leaf and needle fragments, fungi, algae and so on.

Records in polar ice layers²⁶ indicate that major volcanic eruptions often result in massive outputs which dominate the global atmospheric cycle of trace metals. My global inventory suggests that volcanic emanations can account for 40–50% of the Cd and Hg and 20–40% of the As, Cr, Cu, Ni, Pb and Sb emitted annually (Table 1). The data for volcanogenic sources (Table 1) are consistent with those reported recently by Lambert *et al.*¹⁵, who used the release of ²¹⁰Pb to normalize the SO₂ and trace-metal outputs. Patterson and Settle⁵, however, estimated the annual volcanic flux of Pb to be only $0.5 \times 10^{12} \text{ g}$ (corrected from the value of $1.2 \times 10^{12} \text{ g}$ published by these authors), a

value which is lower than the median value but which falls within the range shown in Table 1. The sulphur flux of (15–50) $\times 10^{12} \text{ g}$ used here presumably includes sulphur (and hence is accompanied by trace metals) from both eruptive and non-eruptive activity; fumarolic sulphur emission alone has been estimated to be (15–27) $\times 10^{12} \text{ g yr}^{-1}$ (ref. 27).

With the exception of As (14%) and Sb (16%), seasalt aerosols seem to account for <10% of atmospheric trace metals from natural sources. The seasalt emission rates reported by Wiesel *et al.*²⁸ generally fall towards the upper end of the ranges in Table 1. At present, it is impossible to differentiate between biogenic and (seasalt) spray origins of trace metals in the marine atmosphere, although my inventory tends to favour biological processes as the main contributor. Soil-derived dusts can account for over 50% of the total Cr, Mn and V emissions, as expected. This source may also account for 20–30% of the Cu, Mo, Ni, Pb, Sb and Zn, but <10% of the Hg and Se released annually to the atmosphere. By ignoring biogenic outputs, previous estimates have generally overestimated the importance of resuspended soil particles in the atmospheric trace-metal budget.

The data in Table 1 pertain to emissions in recent years and therefore include some recycled anthropogenic metals. For some metals, the emission rates in prehistoric times were conceivably lower than those calculated today. Biological processing of trace metals dispersed in the environment by human activities is well documented (see, for example, refs 14, 24, 31, 33, 39). It is impossible, however, to use the available information to quantify the contribution of recycled material to the total flux of any of

TABLE 2 Natural versus anthropogenic emissions of trace metals to the atmosphere in 1983

Trace metal	Anthropogenic source	Natural source	Total emission	Natural/total emissions*
As	19 (12–26)	12 (0.86–23)	31 (13–49)	0.39
Cd	7.6 (3.1–12)	1.3 (0.15–2.6)	8.9 (3.2–15)	0.15
Cr	30 (7.3–54)	44 (4.5–83)	74 (12–137)	0.59
Cu	35 (20–51)	28 (2.3–54)	63 (22–105)	0.44
Hg	3.6 (0.91–6.2)	2.5 (0.10–4.9)	6.1 (1.0–11)	0.41
Mn	38 (11–66)	317 (52–582)	355 (63–648)	0.89
Mo	3.3 (0.79–5.4)	3.0 (0.14–5.8)	6.3 (0.93–11)	0.48
Ni	56 (24–87)	30 (3.0–57)	86 (27–144)	0.35
Pb	332 (289–376)	12 (0.97–23)	344 (290–399)	0.04
Sb	3.5 (1.5–5.5)	2.4 (0.07–4.7)	5.9 (1.6–10)	0.41
Se	6.3 (3.0–9.7)	9.3 (0.66–18)	16 (2.5–24)	0.58
V	86 (30–142)	28 (1.6–54)	114 (32–220)	0.25
Zn	132 (70–194)	45 (4.0–86)	177 (74–280)	0.34

All figures in units of 10^9 g yr^{-1} . Emissions from anthropogenic sources are from Nriagu and Pacyna¹; these are median values, with the ranges in estimated emissions given in parentheses.

* Median values only.

the trace metals and metalloids.

The total outputs listed in Table 1 do not agree with the inventories of natural source emissions published by other authors^{2,4,6}. Because of refinement of the biogenic component, the data given here are generally higher than those given in my previous report²⁹. In spite of discrepancies in source intensities, the calculated As flux is in good agreement with the value of $7.8 \times 10^9 \text{ g yr}^{-1}$ reported by Walsh *et al.*², but is less than half of that reported more recently by Chivers and Peterson³⁰. The total annual Hg flux of $(4\text{--}5) \times 10^9 \text{ g}$ reported by Fitzgerald³¹ and the Se flux of $(6\text{--}13) \times 10^9 \text{ g}$ reported by Mosher and Duce¹⁴ are in good agreement with the inventories for these two elements in Table 1.

A comparison of the median values of the worldwide emissions of trace metals from natural and anthropogenic sources suggests that industrial emissions of Pb, Cd and Zn exceed the flux from natural sources by factors of 18, 5 and 3 respectively (see Table 2). Industrial discharges are apparently exercising a profound influence on the atmospheric cycles of these three toxic metals. For As, Hg, Ni, Sb and V, the contributions from anthropogenic sources may exceed those from natural sources by 100–200%. Thus it seems that mankind has become the key agent in the global atmospheric cycle of the toxic metals and metalloids. □

Received 16 August 1988; accepted 11 January 1989.

- Nriagu, J. O. & Pacyna, J. M. *Nature* **333**, 134–139 (1988).
- Pacyna, J. M. *Adv. envir. Sci. Technol.* **17**, 33–52 (1986).
- Walsh, P. R., Duce, R. A. & Fasching, J. L. *J. geophys. Res.* **84**, 1719–1726 (1979).
- Jaworowski, Z., Byssik, M. & Kownacka, L. *Geochim. cosmochim. Acta* **45**, 2185–2199 (1981).
- Patterson, C. C. & Settle, D. M. *Geochim. cosmochim. Acta* **51**, 675–681 (1987).
- Lantzy, R. J. & Mackenzie, F. T. *Geochim. cosmochim. Acta* **43**, 511–523 (1979).
- Duce, R. A. *et al. Rev. Geophys.* **21**, 925–952 (1983).
- Zenchelsky, S. & Youssefi, M. *Rev. Geophys.* **17**, 459–463 (1979).
- Talbot, R. W., Andreae, M. O., Andreae, T. W. & Harris, R. C. *J. geophys. Res.* **93**, 1499–1508 (1988).
- Artaxo, P., Storms, H., Bruynseels, F. & Grieken, R. V. *J. geophys. Res.* **93**, 1605–1615 (1988).
- Orsini, C., Artaxo, P. & Tabacnick, M. *Atmos. Envir.* **16**, 2177–2181 (1982).
- Zoller, W. H. in *Changing Metals Cycles and Human Health* (ed. Nriagu, J. O.) 27–41 (Springer, Berlin, 1983).
- Buat-Menard, P. & Arnold, M. *Geophys. Res. Lett.* **5**, 245–248 (1978).
- Mosher, B. W. & Duce, R. A. *J. geophys. Res.* **92**, 13289–13298 (1987).
- Lambert, G., Le Cloarec, M. F. & Pennisi, M. *Geochim. cosmochim. Acta* **52**, 39–42 (1988).
- Berresheim, H. & Jaeschke, W. *J. geophys. Res.* **88**, 3732–3740 (1983).
- Olmez, I., Finnegan, D. L. & Zoller, W. H. *J. geophys. Res.* **91**, 653–663 (1986).
- Lamb, B., Guenther, A., Guy, D. & Westberg, H. *Atmos. Envir.* **21**, 1695–1705 (1987).
- Zimmerman, P. R., Greenberg, J. P. & Westberg, C. E. *J. geophys. Res.* **93**, 1407–1416 (1988).
- Rasmussen, R. A. & Khalil, M. A. K. *J. geophys. Res.* **93**, 1417–1421 (1988).
- Martell, E. A. & Smith, R. M. *Critical Stability Constants* (Plenum, New York, 1975–77).
- Piotrowicz, S. R., Ray, B. J., Hoffman, G. L. & Duce, R. A. *J. geophys. Res.* **77**, 5243–5254 (1972).
- Schmidt, J. A. & Andren, A. W. *Adv. envir. Sci. Technol.* **14**, 81–103 (1984).
- Toxic Metals in the Atmosphere* (eds Nriagu, J. O. & Davidson) (Wiley, New York, 1986).
- Lindquist, O. & Rhode, H. *Tellus* **37B**, 136–159 (1985).
- Boutroun, C. F. & Patterson, C. C. *J. geophys. Res.* **92**, 8454–8464 (1986).
- Lien, A. Yu. in *The Global Biogeochemical Sulfur Cycle* (eds Ivanov, M. V. & Freney, J. R.) 95–127 (Wiley, New York, 1983).
- Wiesel, C. P., Duce, R. A., Fasching, J. L. & Heaton, R. W. *J. geophys. Res.* **89**, 11607–11617 (1984).

- Nriagu, J. O. *Nature* **379**, 409–411 (1979).
- Chivers, D. C. & Peterson, P. J. in *Lead, Mercury, Cadmium and Arsenic in the Environment* (eds Hutchinson, T. C. & Meema, K. M.) 279–301 (Wiley, New York, 1987).
- Fitzgerald, W. F. in *The Role of Air-Sea Exchange in Geochemical Cycling* (ed. Buat-Menard, P.) 363–408 (Reidel, Dordrecht, 1986).
- Prospero, J. M. *et al. Rev. Phys.* **21**, 1607–1629 (1983).
- Bowen, H. J. M. *Environmental Chemistry of the Elements* (Academic, London, 1979).
- Ure, A. M. & Berrow, M. L. in *Environmental Chemistry* (ed. Bowen, H. J. M.) Vol. 2, 94–205 (Royal Soc. Chem., London, 1982).
- Monahan, E. C. in *The Role of Sea-Air Exchange in Geochemical Cycling* (ed. Buat-Menard, P.) 129–163 (Reidel, Dordrecht, 1986).
- Bruland, K. W. & Frank, R. P. in *Trace Metals in Sea Water* (eds Wong, C. S., Boyle, E., Bruland, K. W., Burton, D. & Goldberg, E.) 415–426 (Plenum, New York, 1983).
- Boyle, E. A. & Huested, S. S. in *Trace Metals in Sea Water* (eds Wong, C. S., Boyle, E., Bruland, K. W., Burton, D. & Goldberg, E.) 379–394 (Plenum, New York, 1983).
- Gill, G. A. & Fitzgerald, W. F. *Geochim. cosmochim. Acta* **52**, 1719–1728 (1988).
- Suzuki, Y. & Sugimura, Y. in *Marine and Estuarine Geochemistry* (eds Sigleo, A. C. & Hattori, A.) 259–273 (Lewis, Chelsea, Ann. Arbor, 1985).
- Jeandel, C., Caisso, M. & Minster, J. F. *Mar. Chem.* **21**, 51–74 (1987).
- Jickells, T. D. & Burton, J. D. *Mar. Chem.* **23**, 131–144 (1988).
- Duce, R. A. *et al. in Marine Pollutant Transfer* (eds Windon, H. & Duce, R. A.) 77–119 (D. C. Heath, Lexington, Mass., 1976).
- Heaton, R. W. thesis, Univ. of Rhode Island (1986).
- Cachier, H., Buat-Menard, P. & Fontugne, M. *Tellus* **38B**, 161–177 (1986).
- Houghton, R. A., Schlesinger, W. H., Brown, S. & Richards, J. F. in *Atmospheric Carbon Dioxide and the Global Carbon Cycle* (ed. Trabalka, J. R.) 114–139 (Office of Energy Research, US Dept of Energy, Washington, D. C., Report no. DOE/ER-0239, 1985).
- Andreae, M. O. *et al. J. geophys. Res.* **93**, 1509–1527 (1988).
- Ajta, G. L., Ketner, P. & Duvigneaud, P. in *Global Carbon Cycle* (eds Bolin, B., Degens, E. T., Kempe, S. & Ketner, P.) 129–182 (Wiley, New York, 1979).
- Curtin, G. C., King, H. D. & Mosier, E. L. *J. geochem. Explor.* **3**, 245–263 (1974).
- Cofer, W. R. *et al. J. geophys. Res.* **93**, 5207–5212 (1988).
- Brown, S. & Hugo, A. E. *Science* **223**, 1290–1293 (1984).
- Andreae, M. O. & Raemondck, H. *Science* **221**, 744–747 (1983).

Studies of static magnetic order in electron-superconductors and their parent compounds

G. M. Luke*, B. J. Sternlieb*, Y. J. Uemura*, J. H. Brewer†, R. Kadono†, R. F. Kiefl†, S. R. Kreitzman†, T. M. Riseman†, J. Gopalakrishnan‡, A. W. Sleight‡, M. A. Subramanian‡, S. Uchida§, H. Takagi§ & Y. Tokura§

* Department of Physics, Columbia University, New York, New York 10027, USA

† Department of Physics, University of British Columbia, Vancouver, British Columbia, Canada V6T 2A6

‡ E. I. DuPont de Nemours & Co., Experimental Station, Wilmington, Delaware 19880-0262, USA

§ Engineering Research Institute and Department of Physics, University of Tokyo, Tokyo 113, Japan

UNTIL recently, all known high- T_c superconductors involved conduction by holes; the doping of holes into the parent compound, for example by the substitution of Sr^{2+} or Ba^{2+} for La^{3+} in $\text{La}_2\text{CuO}_{4-y}$, results in a metallic state whose electronic ground state is superconducting. Recently, Tokura *et al.*¹ have discovered a series of superconducting copper oxides of general formula $\text{Ln}_{2-x}\text{Ce}_x\text{CuO}_{4-y}$ (where Ln represents Pr, Nd, Sm), with transition temperatures as high as $T_c \approx 24 \text{ K}$, in which the charge carriers are electrons. These compounds have the $\text{Nd}_2\text{CuO}_{4-y}$ structure, which is similar to that of $\text{La}_2\text{CuO}_{4-y}$ except for a different arrangement of the oxygen atoms that results in the absence of Cu–O octahedra. Superconductivity is achieved by doping Ce^{4+} for Ln^{3+} and making the samples oxygen-deficient. Here we show, using muon spin rotation, that the parent compounds $\text{Ln}_2\text{CuO}_{4-y}$ of these electron-superconductors exhibit static magnetic order below 300 K, which is similar to that observed in the parent compounds of hole-superconductors. In addition, fairly random static magnetic order has been observed in the non-superconducting material $\text{Nd}_{1.90}\text{Ce}_{0.10}\text{CuO}_{4-y}$ below about 240 K, but no detectable static magnetic order was observed above 4 K in a superconducting sample $\text{Nd}_{1.8}\text{Ce}_{0.16}\text{CuO}_{4-y}$ (with a T_c of 24 K). Hence, electron

doping seems not to fully destroy the magnetic order until the material becomes superconducting. The observation of static magnetic order in these new systems demonstrates that the electronic properties of these copper oxide planar systems are at least partially symmetric with respect to electron and hole doping.

The technique of positive muon spin rotation² (μ SR) has been applied extensively to the study of high- T_c superconductors and related materials³⁻⁸. In solids, the positive muon serves as a sensitive probe of the internal magnetic fields at interstitial sites(s), where the muon resides. In an antiferromagnet one observes one or more discrete oscillation frequencies in zero applied field⁹⁻¹¹, which are independent of the relative orientation of the local field with respect to the initial polarization and are thus the same for a powder and a crystal. In a disordered magnetic system such as a spin glass, in which there is a broad distribution of local-field magnitudes at the muon site(s), the coherence of the muon spins is quickly lost, and this appears as a rapid depolarization of the muon polarization¹².

The application of a weak external transverse magnetic field allows one to determine the volume fraction of the sample that

is magnetically ordered. Muons that reside in sites at which there is an appreciable internal magnetic field (~ 30 G or greater) are quickly depolarized, whereas those in sites with no local ordered field precess in the applied external field¹¹. Thus the fraction of the muon polarization that precesses coherently corresponds to the volume fraction of the sample that has no appreciable static moment.

The pure parent compounds were prepared at DuPont Experimental Station by solid-state reactions of the rare-earth oxides (Nd_2O_3 , Pr_2O_3 , Sm_2O_3) with CuO at $1,000$ – $1,100^\circ\text{C}$, as described elsewhere (J. Gopalkrishnan *et al.*, preprint). X-ray diffraction indicated that the samples were substantially single-phase; powder neutron diffraction revealed the crystal structure to be that of Nd_2CuO_4 (described in ref. 1).

The experiments were performed at the M20 surface muon channel at TRIUMF which provides a beam of $\sim 100\%$ spin-polarized positive muons. Figure 1 shows the zero-field muon polarization at low temperature in $\text{Nd}_2\text{CuO}_{4-y}$, $\text{Pr}_2\text{CuO}_{4-y}$ and $\text{Sm}_2\text{CuO}_{4-y}$. Over much of the temperature range the coherence of the muon spins is lost after a few periods of oscillation. This

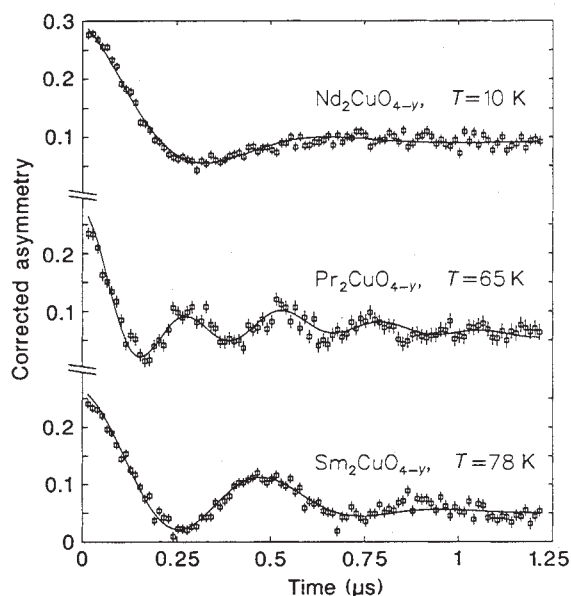


FIG. 1 Zero-field μ SR spectra for $\text{Nd}_2\text{CuO}_{4-y}$, $\text{Pr}_2\text{CuO}_{4-y}$ and $\text{Sm}_2\text{CuO}_{4-y}$ below the ordering temperature.

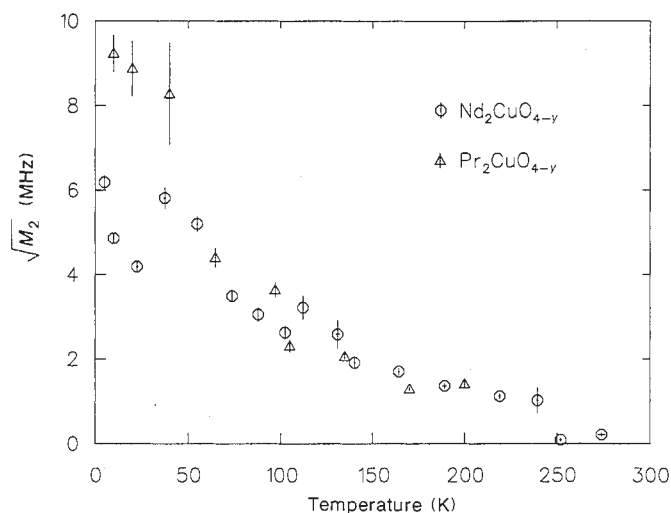


FIG. 3 Temperature dependence of the square root of the muon spin relaxation second moment, measured in zero magnetic field for $\text{Nd}_2\text{CuO}_{4-y}$ and $\text{Pr}_2\text{CuO}_{4-y}$.

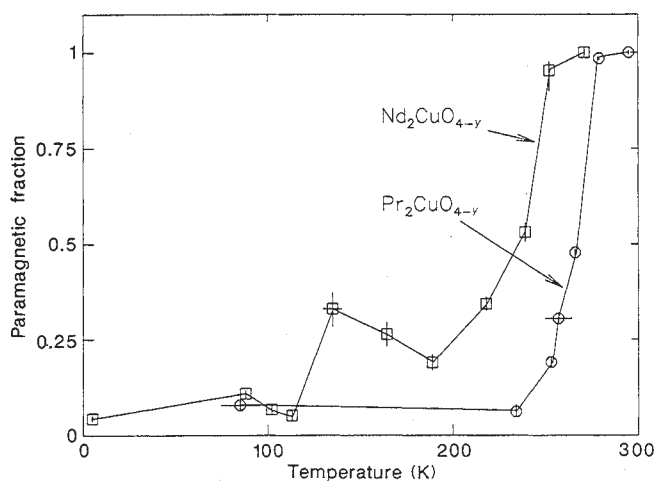


FIG. 2 Temperature dependence of the disordered (paramagnetic) volume fraction in $\text{Nd}_2\text{CuO}_{4-y}$ and $\text{Pr}_2\text{CuO}_{4-y}$.

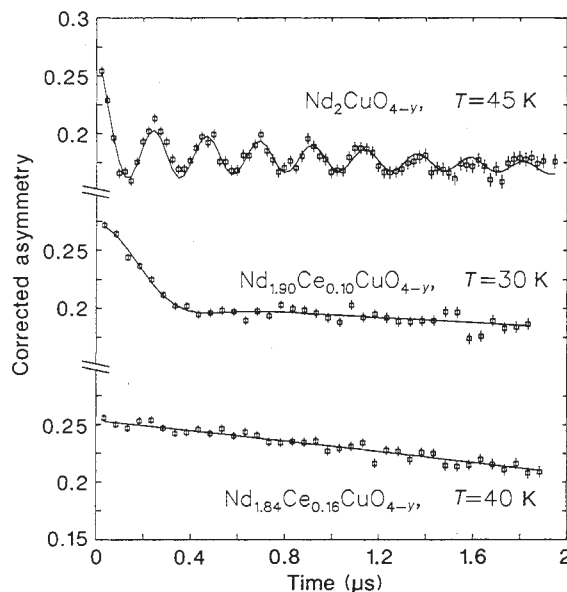


FIG. 4 Zero-field μ SR spectra for reduced samples of non-superconducting $\text{Nd}_2\text{CuO}_{4-y}$ and $\text{Nd}_{1.90}\text{Ce}_{0.10}\text{CuO}_{4-y}$, and superconducting $\text{Nd}_{1.84}\text{Ce}_{0.16}\text{CuO}_{4-y}$.

indicates that the width of the local distribution at the muon site(s) is of the same order of magnitude as the average field. At some temperatures, however, longer-lived precession is clearly visible, indicating that at least short-range magnetic order (rather than mere spin freezing) does occur in these materials. Figure 2 shows the paramagnetic or non-ordered sample volume fraction as a function of temperature for $\text{Nd}_2\text{CuO}_{4-y}$ and $\text{Pr}_2\text{CuO}_{4-y}$.

We have parameterized the muon relaxation function as $\exp(-\frac{1}{2}\sigma^2 t^2) \cos(\omega t)$, corresponding to a gaussian distribution of local fields centred about ω/γ_μ . The fitted frequency ω and the relaxation rate σ are highly correlated when there is such rapid relaxation, so that we then extract the initial second moment of the spin relaxation, given by $M_2 = \sigma^2 + \omega^2$. The square root of the second moment is plotted as a function of temperature in Fig. 3 for $\text{Nd}_2\text{CuO}_{4-y}$ and $\text{Pr}_2\text{CuO}_{4-y}$. The low-temperature values of M_2 are consistent with values reported for $\text{La}_2\text{CuO}_{4-y}$ (refs 9, 10). The similarity in size of the ordered moment leads us to conclude that it is the copper-atom spins that are ordered (as in $\text{La}_2\text{CuO}_{4-y}$), because if the rare-earth atoms were ordered, the ordered moment would be much larger. In a neutron scattering experiment on an essentially similar sample of $\text{Pr}_2\text{CuO}_{4-y}$ below $T = 300$ K, Cox *et al.* (personal communication) have observed Bragg scattering that is consistent with antiferromagnetic ordering of copper spins, and an earlier electron paramagnetic resonance study¹⁵ found indications of antiferromagnetic order of copper spins in the $\text{La}_2\text{CuO}_{4-y}$ materials.

The temperature at which the spins freeze (that is, at which a large static component of the local moment develops) is ~ 300 K for all three compounds. In each of these materials, however, the degree of randomness in the magnetic order exhibits a complex temperature dependence, which will be discussed fully elsewhere.

To study the effects of carrier doping on the magnetic behaviour of these materials, we studied three specimens in the series $\text{Nd}_{2-x}\text{Ce}_x\text{CuO}_{4-y}$: the undoped parent compound $\text{Nd}_2\text{CuO}_{4-y}$, a partially doped, non-superconducting sample $\text{Nd}_{1.90}\text{Ce}_{0.10}\text{CuO}_{4-y}$ and a more fully doped superconducting sample $\text{Nd}_{1.84}\text{Ce}_{0.16}\text{CuO}_{4-y}$, for which $T_c = 24$ K. All three samples, which were prepared at the University of Tokyo¹, were reduced at 900°C under an Ar/O_2 atmosphere with a low oxygen partial pressure, making them oxygen-deficient ($y > 0$). Representative μSR spectra in zero applied field are shown in Fig. 4.

In both non-superconducting samples, static internal fields were observed, indicating some sort of magnetic order or spin freezing. The presence of coherent oscillations (Fig. 4) indicates that more uniform magnetic order occurs in the oxygen-deficient $\text{Nd}_2\text{CuO}_{4-y}$ than in the more fully oxygenated sample shown in Fig. 1. The temperature dependence of this uniformity of order is, however, also complex in this sample. The onset temperature of spin-freezing in the $x = 0.10$ Ce-doped sample was reduced to $T_N \approx 240$ K, and the complete sample volume was ordered or frozen by $T = 150$ K. The magnitude of the static moment at low temperature was approximately the same as that of the parent compound $\text{Nd}_2\text{CuO}_{4-y}$. The presence of static magnetic order at such a high temperature in a material with substantial doping contrasts with the $\text{La}_{2-x}\text{Sr}_x\text{CuO}_{4-\delta}$ system: magnetic order is lost above about 5 K in $\text{La}_{1.92}\text{Sr}_{0.08}\text{CuO}_{4-y}$, and at lower temperatures for higher Sr concentrations¹⁴.

The superconducting $\text{Nd}_{1.84}\text{Ce}_{0.16}\text{CuO}_{4-y}$ sample exhibited no detectable static magnetic order above 4 K. The zero-field relaxation increased somewhat below 60 K (Fig. 4) but was still much less than that seen in non-superconducting samples.

In all of the non-superconducting specimens studied, we observe a more random and complicated magnetic behaviour as a function of temperature than the simple antiferromagnetic ordering observed in the $\text{La}_2\text{CuO}_{4-y}$ system^{9,10}. This behaviour may be due to changes in the ordered spin structure with

temperature, resulting in changing magnetic environments for the muon site(s). An orthorhombic distortion in the $\text{La}_2\text{CuO}_{4-y}$ system stabilizes the magnetic structure of the body-centred copper atoms. In the Ce-doped materials, however, the structure remains tetragonal over the full temperature range and thus the body-centred copper atoms do not have a specific energetically preferred orientation. As there are two equivalent possible orientations of the copper-atom moments, a large degree of randomness in the magnetic order might be expected. Neutron-scattering measurements on single-crystal specimens are necessary to resolve this point, because μSR by itself cannot easily determine the ordered spin structure or the range of the order, as the muons are affected mainly by the nuclei of their nearest neighbours.

Thus the magnetic behaviour of the parent compounds of the electron-superconductors is rather similar to that of the corresponding hole-superconductors, in that ordering occurs at roughly the same temperature with an equivalent static moment. In view of the proposed link between magnetic and superconducting properties in these materials, such information may be of great importance in determining how, if at all, the mechanisms for superconductivity differ for the two classes of compound. $\square$

1. Tokura, Y., Takagi, H. & Uchida, S. *Nature* **337**, 345–347 (1989).
2. Schenck, A. *Muon Spin Rotation Spectroscopy: Principles and Applications in Solid State Physics* (Hilger, Bristol and Boston, 1985).
3. Uemura, Y. J. *et al.* *J. de Physique (Paris) Coll.* (in the press).
4. Uemura, Y. J. *J. appl. Phys.* **64**, 6078–6091 (1988).
5. Nishanda, N. *et al.* *Jap. J. appl. Phys.* **26**, L1856–L1858 (1987).
6. Aepli, G. *et al.* *Phys. Rev.* **B35**, 7129–7131 (1987).
7. Gygax, F. N. *et al.* *Europhys. Lett.* **4**, 473–479 (1987).
8. Uemura, Y. J. *et al.* *Nature* **335**, 151–152 (1988).
9. Uemura, Y. J. *et al.* *Phys. Rev. Lett.* **59**, 1045–1048 (1987).
10. Uemura, Y. J. *et al.* *Physica C* **153–155**, 769–770 (1988).
11. Brewer, J. H. *et al.* *Phys. Rev. Lett.* **60**, 1073–1075 (1988).
12. Uemura, Y. J., Yamazaki, T., Harshman, D. R., Senba, M. & Ansaldo, E. J. *Phys. Rev.* **B31**, 546–563 (1985).
13. Saez Puche, R., Norton, M., White, T. R. & Glaunsinger, W. S. *J. Solid State Chem.* **50**, 281–293 (1983).
14. Weidinger, A. *et al.* *Phys. Rev. Lett.* **62**, 102–105 (1989).

Wavelet analysis of turbulence reveals the multifractal nature of the Richardson cascade

F. Argoul*, A. Arnéodo*, G. Grasseau*, Y. Gagne†, E. J. Hopfinger† & U. Frisch‡

* Centre de Recherche Paul Pascal, Laboratoire Propre CNRS, Domaine Universitaire, 33405 Talence Cedex, France

† Institut de Mécanique de Grenoble, Unité Mixte de Recherche CNRS, BP 53X, 38041 Grenoble Cedex, France

‡ Observatoire de Nice, Unité Associée CNRS, BP 139, 06003 Nice Cedex, France

THE central problem of fully developed turbulence is the energy-cascading process. It has resisted all attempts at a full physical understanding or mathematical formulation. The main reasons for this failure are related to the large hierarchy of scales involved, the highly nonlinear character inherent in the Navier–Stokes equations, and the spatial intermittency of the dynamically active regions. Richardson¹ has described the interplay between large and small scales with the words ‘Big whirls have little whirls which feed on their velocity ... and so on to viscosity’, and the phenomenon so described is known as the Richardson cascade. This local interplay also forms the basis of a theory by Kolmogorov². It was later realized that the cascade ought to be intermittent, and Mandelbrot³ has given a fractal description of the intermittency of the fine structure. A particular case that emphasizes the dynamical aspect of the fractal models is the β -model⁴, in which the flux of energy is transferred to only a fixed

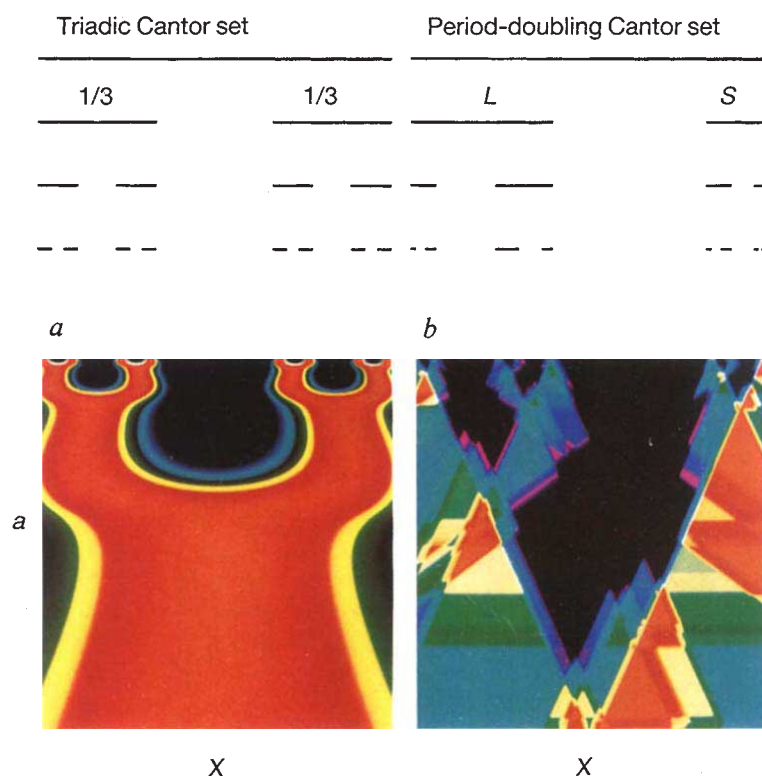


FIG. 1 Wavelet analysis of Cantor sets^{9,15}. Coordinates correspond to space (x) and scale (a); both are linear (small scales at the top). The intensity of the wavelet transform is colour-coded according to the natural light spectrum from black (low intensity) ($T_g \leq 0$) to red ($\max(T_g) > 0$) on each line $a = \text{constant}$. *a*, Uniform triadic Cantor set analysed using a 'Mexican hat' wavelet: $g(x) = (1 - x^2)e^{-x^2/2}$; $x \in [0, 1]$, $a \in [2^{-9}, 2^0]$. *b*, Period-doubling Cantor set analysed using 'French hat' wavelets: $g(x) = 1(|x| < 1)$, $-1/2(1 \leq |x| \leq 3)$, $0(|x| > 3)$; $x \in [-0.40115, \dots, 1]$, $a \in [2^{-10}, 2^{-1}]$. Above each figure we illustrate the process of subdivision leading to each Cantor set.

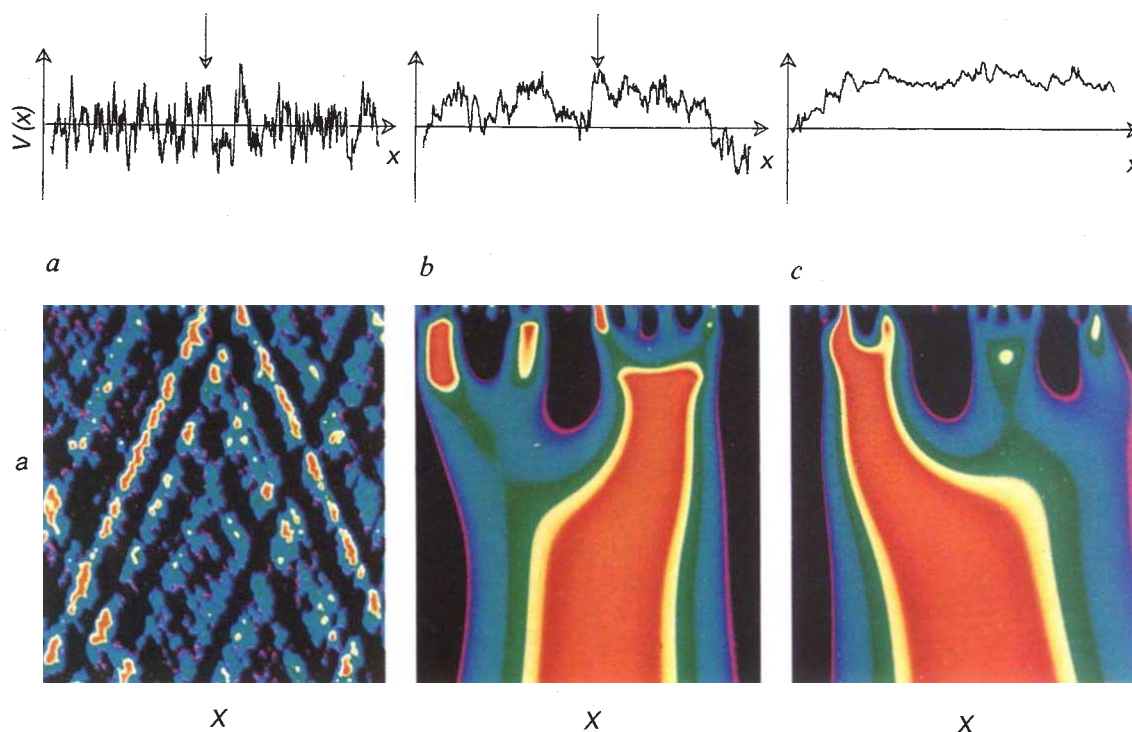


FIG. 2 Wavelet analysis of fully developed turbulence from wind-tunnel data. Coordinates and intensities are as in Fig. 1. The top graphs show the signals being analysed. Panel *a* is a 852-m-long sample, showing the large scales (from $28l_0$ to $l_0/10$) analysed using a 'French top hat' wavelet. In *b* the

'Mexican hat' wavelet was used for magnification $\times 20$ of the central position indicated by the arrow in the top graph of *a*. Panel *c* bears an analogous relation to *b*, again using the 'Mexican hat' wavelet. The successive forkings reveal the fractal nature of the Richardson cascade.

fraction β of the eddies of smaller scales. More recently, a multifractal model of the fine-scale intermittency has been introduced by Parisi and Frisch⁵; this accounts for the more complex cascading process suggested by the experimental data on inertial-range structure functions obtained by Anselmetti *et al.*⁶. These statistical models are insufficient, however, because they cannot give any topological information about the intermittent structures in real space. Here we use the wavelet transform⁷ to analyse the velocity field of wind-tunnel turbulence at very high Reynolds numbers⁸. This 'space-scale' analysis is shown to provide the first visual evidence of the celebrated Richardson cascade¹, and reveals in particular its fractal character³. The results also indicate that the energy-cascading process has remarkable similarities with the deterministic construction rules of non-homogeneous Cantor sets^{9,10}.

Wavelet analysis⁷ is a mathematical technique introduced recently for analysing seismic data and acoustic signals^{11,12}. It provides a two-dimensional unfolding of one-dimensional signals, resolving both the position and the scale as independent variables. The method comprises an expansion of an arbitrary real-valued function $s(x)$ over wavelets that are constructed from a single function g by means of dilatations and translations^{7,13}. The wavelet transform of $s(x)$ with respect to the wavelet g is defined as:

$$T_g(a, x) = w(a) \int g((x-y)/a) s(y) dy \quad a > 0, x \in \mathbb{R}$$

where the weight function $w(a)$ is introduced for visual enhancement.

The wavelet g is a regular function, localized around $x=0$. For a large class of functions $s(x)$, the wavelet transform may be inverted, provided that g is admissible ($\int g(y) dy = 0$) and satisfies other conditions^{7,14}. The wavelet transformation can be regarded as a mathematical microscope⁹, for which position and magnification correspond to x and a^{-1} , respectively, and the performance of the optics is determined by the choice of the analysing wavelet g . Wavelet analysis has become a powerful tool for locating singularities^{9,10}: a singularity of $s(x)$ at x_0 produces a cone-like structure in T_g , pointing towards the point ($a=0, x=x_0$). The nature of the singularities may be inferred from the power-law behaviour of $T_g(a, x=x_0)$ on increasing the magnification a^{-1} .

As emphasized by Arnéodo *et al.*⁹, the wavelet transform assists visualization of self-similar properties of fractal objects. In particular, it illustrates the complexity of the fractal under consideration, revealing the hierarchy that governs the relative positioning of the singularities. In Fig. 1a, the successive pitchfork branchings observed in $T_g(a, x)$ as a result of increasing the magnification a^{-1} illustrate how the uniform (single-scale) triadic Cantor set is constructed^{9,10}. The unit interval $[0, 1]$ is initially divided into three sub-intervals, each of length $l=1/3$. The middle interval is removed and the same process is repeated on each of the two remaining sub-intervals, and so on *ad infinitum*. In Fig. 1b, the same analysis is performed for the multifractal (two-scale) Cantor set that arises at the onset of chaos through the period-doubling cascade scenario^{9,15}. The pitchfork branchings observed on increasing the magnification a^{-1} are no longer symmetric, and occur not in pairs but successively, as the signature of two unequal scaling factors, one large and one small ($L(\alpha_{PD}^{-1}) \approx 1/2.5029 \dots$ and $S(\alpha_{PD}^{-2})$ respectively). Moreover, the internal self-similarity of T_g (which is central to the renormalization-group study of the period-doubling transition to chaos) illustrates the basic construction rule of the period-doubling Cantor set: from one stage to the next, the relative positions of the large and small sub-intervals is unchanged when dividing a small interval, but is exchanged when dividing a large interval.

The fractal structures proposed in fully developed turbulence are real-space structures, in contrast to fractal attractors, which reside in phase space. Thus the wavelet transform can be applied

directly to turbulent velocity fields obtained from experiment or from numerical simulations¹⁶. Here we report the first such analysis performed on the velocity data from high-Reynolds-number, three-dimensional turbulence. The data were obtained in the large wind tunnel S1 of ONERA in Modane, France. A hot wire probe was located on the axis of the return section which was 24 m in diameter and 150 m in length. The turbulence level, expressed by the ratio of the r.m.s. value of the longitudinal velocity fluctuations to the mean velocity, was less than 7%, which implies that time variations are equivalent to spatial variations (the Taylor hypothesis). The scale ratio of the velocity signal shown in Fig. 2 is about 4×10^4 , giving 15 m for the integral scale l_0 and 0.35 mm for the dissipation scale l_d . The Reynolds number based on the Taylor microscale is $R_\lambda = 2,720$. This turbulence has an inertial range of at least two decades^{8,17}.

A signal of total length 852 m was analysed hierarchically. Figure 2a corresponds to the total span and to a range of scales (that is, of parameter a) from $28l_0$ to $l_0/10$. Figure 2b shows the region around the arrow in Fig. 2a, with position and scale variables magnified by a factor of 20, and Fig. 2c is related to Fig. 2b in the same way. A 'French top hat' piecewise constant analysing wavelet was used in the case of Fig. 2a, emphasizing the cone-like structures. A 'Mexican hat' wavelet was used in Figs 2b and c, because it provides a (marginally) better visualization of the Richardson cascade¹. Only positive values of the wavelet transform are shown. In the figures the weight function $w(a)$ was chosen such that the maximum of $T_g(a, x)$ on each $a = \text{const.}$ line is independent of a .

Large-scale eddies are revealed in Fig. 2a by the superimposed cone-like patterns of different intensities, indicating structures of scale $a \sim l_0$. These large-scale structures seem randomly distributed in space and are fairly space-filling. Very different small-scale patterns are displayed in Figs 2b and c. These are created by a process in which typically an eddy divides asymmetrically (in intensity) into two smaller eddies, and this subdivision is repeated about eight to ten times before eddies reach a scale at which they are readily dissipated. These successive forkings produce a fractal structure, in which small-scale activity becomes increasingly sparse. The ratio between the scales of successive generations can take different values, which is another indication that a multifractal description is appropriate.

In summary, we have used wavelet analysis, which is already known to be a powerful tool for analysing multifractal attractors^{9,15}, to explore the spatial structure of fully developed turbulence. Our emphasis has been to identify new qualitative features. Similar techniques can be used to reveal fractal clustering of galaxies on very large scales (A. Bijaoui & E. Slezak, unpublished work). $\square$

Received 5 December 1988; accepted 6 January 1989.

1. Richardson, L. F. *Weather Prediction by Numerical Process* (Cambridge Univ. Press, 1922).
2. Kolmogorov, A. N. *C. r. Acad. Sci. USSR* **30**, 301–358 (1941).
3. Mandelbrot, B. B. *J. Fluid Mech.* **62**, 331–358 (1974).
4. Frisch, U., Sulem, P. L. & Nelkin, M. *J. Fluid Mech.* **87**, 719–736 (1978).
5. Parisi, G. & Frisch, U. in *Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics* (eds Ghil, M., Benzi, R. & Parisi, G.) 71–88 (North-Holland, Amsterdam, 1985).
6. Anselmetti, F., Gagne, Y., Hopfinger, E. J. & Antonia, R. A. *J. Fluid Mech.* **140**, 63–89 (1984).
7. Grossmann, A. & Morlet, J. in *Mathematics and Physics, Lecture on Recent Results* (ed. Streit, L.) (World Scientific, Singapore, 1987).
8. Gagne, Y. thesis, Univ. of Grenoble (1987).
9. Arnéodo, A., Grasseau, G. & Holschneider, M. *Phys. Rev. Lett.* **61**, 2281–2287 (1988); Arnéodo, A., Grasseau, G. & Holschneider, M. in *Wavelets* (eds Combes, J. M., Grossmann, A. & Tchamitchian, P.) (Springer, Berlin, in the press).
10. Holschneider, M. *J. stat. Phys.* **50**, 963–993 (1988).
11. Goupillaud, P., Grossmann, A. & Morlet, J. *Geoelectrochim. Acta* **23**, 85–102 (1984).
12. Kronland-Martinet, R., Morlet, J. & Grossmann, A. *Int. J. Pattern Recognition and Artificial Intelligence*, spec. iss. on *Expert Systems and Pattern Analysis* (in the press).
13. Daubechies, I., Grossmann, A. & Meyer, Y. *J. math. Phys.* **27**, 1271–1283 (1986).
14. Grossmann, A., Morlet, J. & Paul, T. *J. math. Phys.* **26**, 2473–2479 (1985).
15. Arnéodo, A., Argoul, F., Elzgaray, J. & Grasseau, G. in *Nonlinear Dynamics* (ed. Turchetti, G.) (World Scientific, Singapore, in the press).
16. Farge, M. & Rabreau, G. *C. r. Acad. Sci. Paris* **307**, Ser. II, 1479–1488 (1988).
17. Gagne, Y., Hopfinger, E. J. & Frisch, U. in *New Trends in Nonlinear Dynamics and Pattern Forming Phenomena: The Geometry of Non-Equilibrium* (eds Huerre, P. & Coulet, P.) (Plenum, New-York, 1988).

ACKNOWLEDGEMENTS. We acknowledge the experimental facilities provided by ONERA and financial support from the EEC.

The 1988 US drought linked to anomalous sea surface temperature

T. N. Palmer & Č. Branković

European Centre for Medium Range Weather Forecasts,
Shinfield Park, Reading, RG2 9AX, UK

THE 1988 drought in the United States has been widely reported, not least as an indicator of the reality of the 'greenhouse effect'—see for example ref. 1. On the other hand, drought is a naturally occurring phenomenon. Here we have studied 30-day forecasts of the atmospheric flow over the United States using a complex numerical weather-prediction model initialized with data from May 1987 and May 1988. In addition, an experiment with 1988 initial conditions and 1987 sea surface temperatures was made. The results indicate that much of the difference between the 1987 and 1988 forecasts was associated with interannual variability in sea surface temperature. These results provide complementary evidence to that in ref. 2, which suggested that the US drought was linked to anomalous oceanic conditions in the tropical Pacific.

The El Niño event in the tropical Pacific ocean is now recognized as part of the natural variability of the coupled ocean-atmosphere system³. At its height, the El Niño event is associated with an anomalous warming of the surface waters of the east and central tropical Pacific. There have been many studies suggesting that in the northern winter, there is a link between this warming and unusual weather conditions in the extratropics⁴. During 1987 there was a moderate El Niño event. In 1988, however, conditions in the tropical east Pacific began to swing towards the opposite, 'anti El Niño' or 'La Niña', phase of the event where sea surface temperatures (SSTs) became anomalously negative. Little is known about the possible impact of La Niña in the extratropics.

The difference between SSTs averaged for the 15 days preceding 24 May 1987 and 22 May 1988 is shown in Fig. 1a. The relative warmth of the equatorial east Pacific in 1987 relative to 1988 (over 3 K), is clearly shown. To the north and south of this equatorial band, SST differences were reversed. It should be pointed out, however, that there were significant SST differences in other parts of the globe. As shown, differences of 3 K were also attained in the subtropical north Pacific, and differ-

ences of 2 K were observed in the Atlantic and Indian oceans. Whether these are causally linked to the El Niño cycle is unknown at present.

Figure 1b shows the observed time-average difference in the height of the atmospheric 300-mbar pressure surface between 25 May–23 June 1987 and 23 May–21 June 1988. Over much of North America the pressure surface was lower in 1987 than 1988. This is consistent with the fact that, during the 1988 drought, the jet stream and associated rain-producing weather systems were displaced northward of their usual positions by anomalous high pressure. Relative to 1988, the height of the 300-mbar surface was also depressed over a region extending from Alaska down to the subtropical north-east Pacific. Note also that there is a rather strong negative depression over the north-west Atlantic.

Based on observations of the anomalous height fields for 1988, and using theoretical concepts of Rossby-wave radiation from a tropical heat source, and a relatively simple numerical model, Trenberth *et al.*² suggested that the 1988 US drought and anomalous tropical Pacific SSTs were causally related, and inferred that the 1988 US drought was part of normal climatic variability.

The European Centre for Medium Range Weather Forecasts (ECMWF) has developed a complex and comprehensive global numerical weather-prediction model⁵ which is integrated daily to give weather guidance up to ten days ahead. On an experimental basis over the past four years, the model has also been integrated to 30 days, twice a month, from initial data separated by 24 h. In general, the day-to-day variations of weather are not predictable beyond about two weeks into the future. However, from time to time, the 30-day extended range forecasts do show impressive skill (that is, the forecast and observations correlate well).

Figure 2 shows some 30-day mean difference fields from extended-range forecasts initialized on 24 May 1987 and 22 May 1988. Comparing the forecast 300-mbar height field difference in Fig. 2a with the observed height field difference (Fig. 1b), it can be seen that the difference in model simulations for the two years is quite skilful. In particular, the three depressions of the actual 300-mbar height surface described above, over the Atlantic, North America and the North-east Pacific, all have counterparts in the forecast field. Although the two fields clearly do not correlate perfectly, the verisimilitude of this extended-range forecast difference field is in general above average. Indeed the difference field for two 30-day forecasts, each initialized just

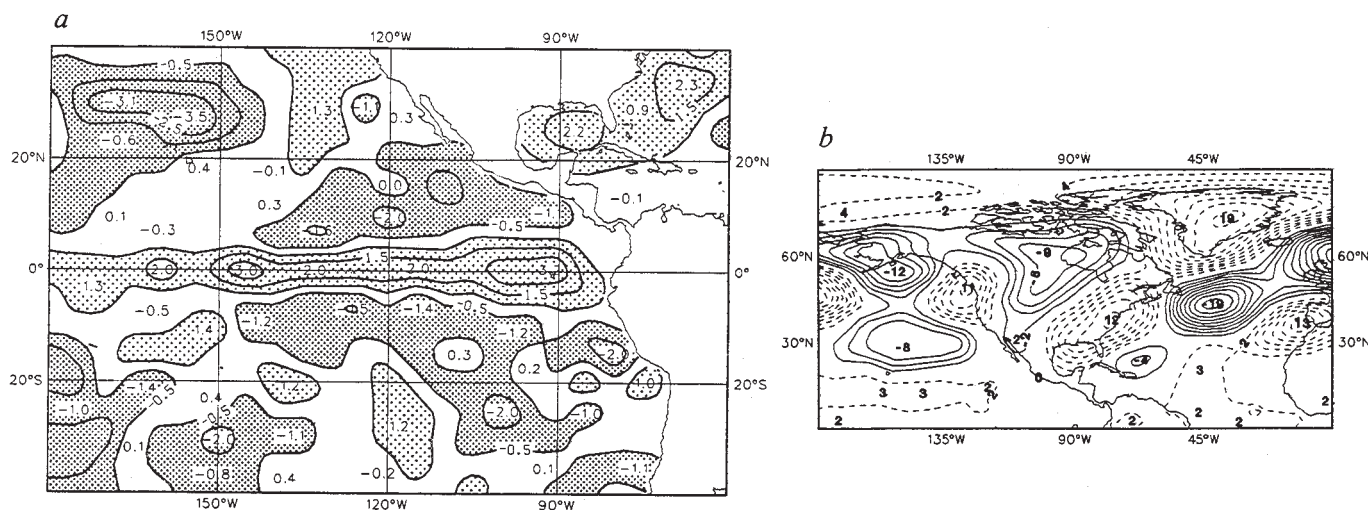


FIG. 1 a, Difference in sea surface temperatures between data averaged for 15 days preceding 24 May 1987, and for 15 days preceding 22 May 1988. Contour interval of 1 K. Areas above 0.5 K shown with coarse stipple; areas below -0.5 K shown with fine stipple. The two SST fields whose difference is shown here were used as oceanic boundary conditions for the model

integrations discussed in the text. Data from operational US National Meteorological Center analyses. b, Time-average difference in the height of the 300-mbar pressure surface between the 30 days following 24 May 1987 and the 30 days following 22 May 1988. Contour interval 20 cm. Positive values are shown with dashed contours.

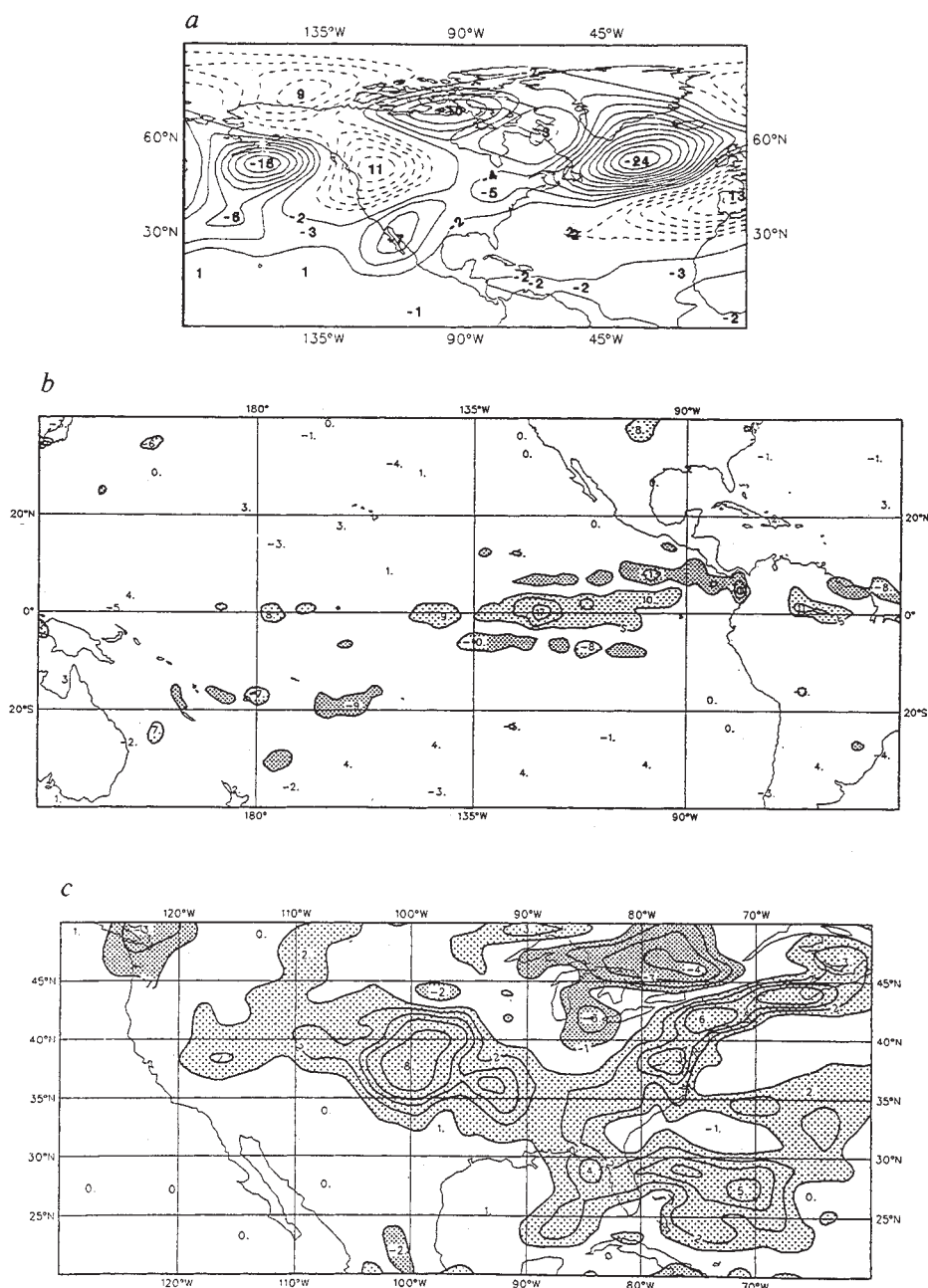


FIG. 2 Thirty-day average difference fields between two forecast integrations of the (T63) ECMWF numerical weather prediction model, the first initialized from data for 24 May 1987, the second from 22 May 1988. *a*, The height of the 300-mbar pressure surface, contours as Fig. 1*b*. *b*, Rainfall over the east tropical Pacific with contour interval of 5 mm per day (0 contour suppressed), coarse stippling above 5 mm per day, fine stippling below -5 mm per day. *c*, Rainfall over North America with contours of ± 6 , ± 4 , ± 3 , ± 2 , ± 1 mm per day; coarse stippling above 1 mm per day, fine stippling below -1 mm per day.

one day earlier than those shown here, was very much poorer. The sensitivity of extended-range forecasts to initial conditions can be expected on fundamental theoretical grounds⁶.

Figure 2*b* shows 30-day mean forecast difference rainfall fields over the tropical Pacific. Over the equatorial east Pacific there is a clear increase in rain in 1987 compared with 1988. To the immediate north and south, there is a relative decrease in 1987. This field is clearly correlated with the SST difference field shown in Fig. 1*a*, and agrees with satellite remote-sensing proxy observations of rainfall².

Figure 2*c* shows 30-day mean forecast difference rainfall fields over North America. Over much of the eastern and central United States, the differences are generally positive and substantial, indicating in some areas a forecast of rainfall for 1988 less than 20% of that for 1987. Of course, these rainfall difference fields are not accurate in detail, but certainly capture much of the essence of the 1988 drought⁷.

The question arises as to whether the difference fields shown in Fig. 2 reflect primarily the fact that the initial conditions between the two integrations were different, or whether the

'boundary conditions' also played an important role. In numerical weather-prediction models (as opposed to many climate models), SSTs are specified and held fixed during an integration, and are therefore part of the boundary conditions for such models. For the ECMWF forecasts, SSTs are obtained as an average of observations for 15 days preceding the initial date. Figure 1*a* shows the difference field of SSTs used in the two forecasts.

To examine the importance of SST differences in accounting for the difference fields in Fig. 2, we have run a third 30-day integration using identical conditions to the 1988 forecast, but with SSTs from the 1987 integration. The difference in 30-day mean 300-mbar height and rainfall between this experiment and the 1988 forecast integration is shown in Fig. 3.

In testing the potential impact of SST anomalies on the atmosphere, it is common to run a single climate timescale integration of several years, or a large number of shorter timescale integrations from a variety of initial conditions. In the latter case, the impact of SSTs will in general depend quite strongly on the particular initial conditions^{8,9}, and it may be

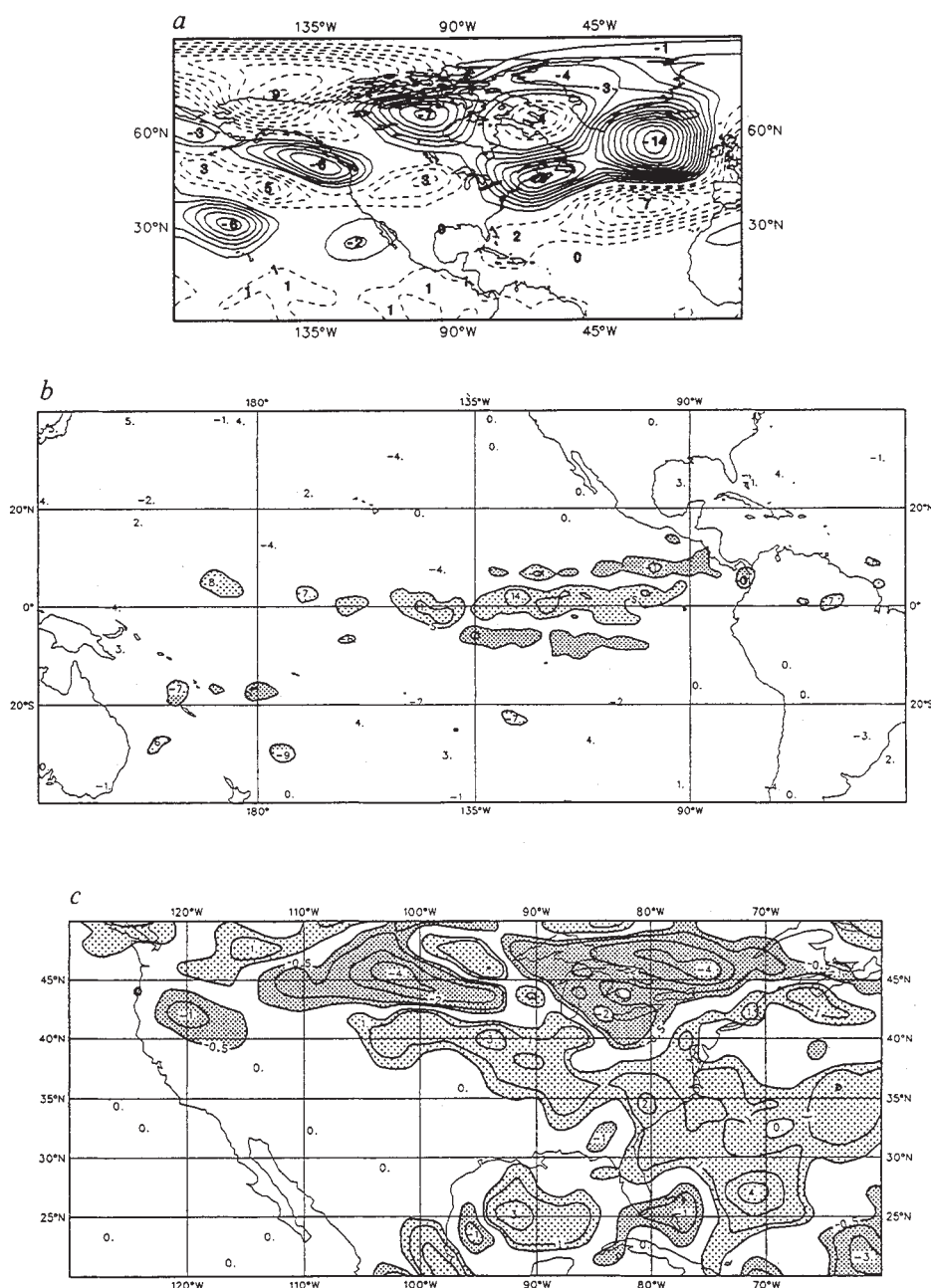


FIG. 3 Thirty-day average difference fields between two integrations of the (T63) ECMWF numerical weather prediction model; the first initialized from data for 22 May 1988 and sea surface temperatures from the 15 days preceding 24 May 1987, the second initialized from the same data but with sea surface temperatures averaged from the 15 days preceding 22 May 1988. *a*, The height of the 300-mbar pressure surface; contour interval 10 cm. *b*, Rainfall over the east tropical Pacific, contours as Fig. 2*b*. *c*, Rainfall over North America with contours of ± 5 , ± 3 , ± 2 , ± 1 , ± 0.5 mm per day; coarse stippling above 0.5 mm per day, fine stippling below -0.5 mm per day.

that one would deduce that the response of the atmosphere to a particular SST anomaly was not strongly significant. If one has the option, however, of verifying the shorter-timescale integrations as forecasts, one finds that many, if not most, show little extended-range forecast skill, although a small (maybe very small) fraction show exceptional extended-range skill. The question then arises as to whether the response of the unskilful forecasts to the imposed SST anomalies should be given the same weight as the response to the skilful forecasts (as would conventionally be the case). We suggest that they should not. This is the reason why we choose to emphasize the response to our skilful forecast, rather than the unskilful forecast initialized one day earlier. (In fact one could broadly think of our results in terms of an ensemble mean of two forecast experiments, weighted *a posteriori* by the extended-range skill of the basic forecast.)

The 300-mbar height difference field (Fig. 3*a*) is clearly correlated with the forecast difference field shown in Fig. 2*a*, although the amplitude of the field is smaller (note the reduced contour interval in Fig. 3*a*). Note, in particular, the depression of the

height of the 300-mbar surface over the Atlantic and over North America. The depression over the North Pacific is certainly not as strong in Fig. 3*a* as in Fig. 2*a*, although the general pattern of wind anomalies that would result from these height field differences corresponds well.

Rainfall differences over the tropical Pacific are shown in Fig. 3*b*. Comparing with Fig. 2*b*, there can be no doubt that the simulated difference between 1987 and 1988 rainfall in that region is primarily the result of the different boundary conditions.

Rainfall differences over North America are shown in Fig. 3*c*. It can be seen that there is a general increase in rainfall with 1987 SSTs to the south of 40° N, and the pattern is well correlated with the forecast difference in Fig. 2*c*. Moreover, there is a decrease in rainfall over the Great Lakes area, consistent with the decrease there in Fig. 2*c*. Again, it should be emphasized that Fig. 3 should be compared with Fig. 2 rather than the observations; shortcomings in the veracity of the forecast fields are reflected in similar shortcomings for the experiment.

These results clearly indicate that the impact of the SSTs was

in the same sense as the difference between the two forecasts. The amplitudes in Fig. 3 are clearly smaller than in Fig. 2, because the impact of the SSTs is not felt to any significant extent in the first few days of the forecast.

Hence we conclude that the anomalous SSTs in 1987 and 1988 were indeed important in accounting for the reduction in rainfall over the United States in the late spring of 1988. We cannot, on the basis of these results, state that La Niña itself (rather than other regions of SST anomaly) was principally important. Further experiments should help to clarify this. However, on the basis of this study and the investigation of Trenberth *et al.*², using very different techniques to those discussed here, tropical Pacific ocean temperatures related to the El Niño/La Niña event cycle appear to be important, and there now seems to be mounting evidence that the 'cause' of the 1988 US drought was primarily associated with phenomena that are part of the natural variability of the climate system. □

Received 14 December 1988; accepted 25 January 1989.

1. Palca, J. *Nature* **334**, 92 (1988).
2. Trenberth, K. E., Branstator, G. W. & Arkin, P. A. *Science* **242**, 1640–1645 (1988).
3. Cane, M. A., Zebiak, S. E. & Dolan, S. C. *Nature* **321**, 827–832 (1986).
4. Rasmusson, E. M. & Wallace, J. M. *Science* **222**, 1195–1202 (1983).
5. Simmons, A. J. *Adv. Geophys.* **29**, 305–338 (1986).
6. Lorenz, E. N. *J. Atmos. Sci.* **20**, 130–141 (1963).
7. *Climate System Monit. Mon. Bull.* Issue No 6 (WMO Secretariat, Geneva 1988).
8. Mansfield, D. A. Q. *J. R. met. Soc.* **112**, 1145–1176 (1986).
9. Tribbia, J. J. & Baumhefner, D. P. *J. Atmos. Sci.* **45**, 2306–2317 (1988).

Organic geochemical evidence for global fires at the Cretaceous/Tertiary boundary

M. I. Venkatesan & J. Dahl

Institute of Geophysics and Planetary Physics,
University of California, Los Angeles, California 90024-1567, USA

MANY hypotheses have been advanced to explain the mass extinction at the Cretaceous/Tertiary (K/T) boundary^{1–3}. Recently, Wolbach *et al.*⁴ suggested that massive forest fires were triggered by the impact of a meteorite, and cite as evidence the presence of elemental carbon (mainly soot) from K/T boundaries⁵. Almost all of the airborne polycyclic aromatic hydrocarbons (PAHs) generated by pyrosynthesis are adsorbed, through hydrogen bonding, on the surface of soot, the particulate fraction from combustion^{6,7}. Although soot itself is a polymer of polybenzenoid radicals, early termination of polymerization leads to enhanced PAH production. Pyrosynthesis of PAHs is thus favoured by a chemically reducing atmosphere. If there were wildfires, a group of high-molecular-weight parent PAHs characteristic of combustion, predominating over their alkyl homologues^{8,11}, should be present in K/T boundary samples known to contain soot^{4,5}. Here we compare K/T samples from New Zealand, Italy and Denmark to those from above and below the boundary, and find enhanced PAH contents and distribution profiles that reflect a pyrolytic origin. The data thus provide the first detailed organic-molecular evidence for the combustion source of organic carbon at the K/T sites.

The organic-rich clay layer (corresponding to S6 to S10 of Schmitz¹²) above the red layer was collected by B. Schmitz by digging into the Fish Clay outcrop at Stevns Klint, Denmark (200 m south of the Højrup Church). Only the unweathered clay was gathered. Woodside Creek boundary clay (upper layer) and two samples, one well into the Tertiary and the other well into the Cretaceous, were obtained from R. R. Brooks (Table 1). The exact location, stratigraphy and other details of depositional environment are available^{13,14}. The red boundary clay, and a sample from above and below the boundary, from Bottaccione Gorge, Gubbio, Italy, were collected by W. Cavazza.

Sample location and trace-element distribution are described in ref. 15. The samples were shipped by air in sealed plastic bags, some first wrapped in aluminium foil.

Chunk samples were quickly sonicated in dichloromethane to remove external contamination. Dry powdered samples were then extracted in dichloromethane and fractionated into aliphatic and aromatic hydrocarbons and other classes of compound according to the methods of Venkatesan *et al.*¹⁶, which involve thin-layer and column chromatography. The aromatic hydrocarbon fraction (containing PAHs) was analysed by gas chromatography and quantified by gas chromatography/mass spectrometry (GC/MS), based on the response factor of an internal standard (dodecylbenzene) and an external standard mixture containing PAH compounds listed under Fig. 1.

Pertinent geochemical data are presented in Table 1. The iridium enrichment at the boundary clays is generally concomitant with the enhanced levels of organic carbon and total PAH content. This is also consistent with a recent observation¹⁷ that the bulk of the iridium is probably organically complexed.

The distribution of PAHs in K/T extracts is illustrated in Fig. 1. The Stevns Klint and Gubbio samples are enriched in high-molecular-weight PAHs relative to the Woodside Creek sample. Phenanthrene is the most dominant and 2-methylfluorene the second most dominant PAH in the Woodside Creek sample. In the Stevns Klint sample, coronene is the most abundant PAH, followed by phenanthrene. Coronene is also the most dominant PAH in the Gubbio sample, followed by benzo(*b+k*)fluoranthene and chrysene/triphenylene. Pyrene and fluoranthene are significant constituents in the Stevns Klint and Gubbio samples. The dominant analogues in both of these samples are the peri-condensed aromatic series, such as pyrene, benzo-pyrene, benzo(*g,h,i*)perylene and coronene. These peri-condensed compounds are formed by the pyrolysis of organic matter¹⁸ such as gasoline¹⁹ and wood²⁰, and are relatively uncommon in petroleum⁶ but have been detected in significant amounts in hydrothermal mounds²¹. Because peri-condensed structures are more reactive than their angularly-condensed analogues¹⁸, the presence of the former in significant amounts in these samples implies that rapid thermal quenching by adsorption of these PAHs in soot contributed to their stability. Also, the abundance of these pyrolytic PAHs should have been high compared with that seen today, the latter representing the residual combustion PAHs after degradation¹⁹.

The absence of peri-condensed PAHs in the Woodside Creek K/T sample could be attributed to one of the following. (1) They were not formed because there were no wildfires. (2) The continent adjacent to this site experienced elevated temperatures for longer intervals relative to the other two sites, with

TABLE 1 Pertinent geochemical data for K/T site samples

Sample	Total % organic carbon	Ir (p.p.b.)†	Total PAH* (p.p.b.)
Woodside Creek, New Zealand			
K/T	0.72	26‡	33
K/T + 250 cm	0.02	NA	5
K/T – 275 cm	0.01	NA	3
Stevns Klint, Denmark			
K/T	1.40	50	30
Gubbio, Italy			
K/T	0.10	5§	75
K/T + 5–20 cm	NA	NA	<<1
K/T – 3–15 cm	NA	NA	<<1

NA, Not available.

* Total PAH summed from naphthalene to 1, 2, 4, 5-dibenzopyrene listed under Fig. 1.

† Parts per 10⁹.

‡ Personal communication from R. R. Brooks.

§ Data from ref. 15.

consequent elimination of the less stable, intermediate¹⁸, peri-condensed PAHs. No literature data is available at present on the stability of peri-condensed PAHs and their burial history. The presence of five-membered alicyclic rings such as 2-methylfluorene, 2,3-benzofluorene and fluoranthene also attest to the pyrolytic origin of PAHs in all three of the sites. These PAHs are commonly present in all pyrolysates from organic matter and, once formed, are not easily converted to peri-condensed aromatics, because they inhibit subsequent enlargement of the aromatic nuclei^{18,22}. Thus the chemical composition of the PAH fraction suggests a pyrolytic source. The Woodside Creek K/T sample, therefore, reflects input of PAHs from combustive sources, possibly from an adjacent land mass that experienced extended periods of high temperatures caused by forest fires, resulting in the destruction of less stable, linear PAHs¹⁸. (3) The Woodside Creek site shows considerable lateral variations in boundary-clay thickness and Ir abundances, which are apparently also reflected in the PAH profile. This may explain the relatively high level (1.4 parts per 10⁹) of coronene reported earlier from a Woodside Creek K/T layer²³.

The dominance of parent PAHs is characteristic of combustion products, whereas alkylated PAHs predominate in bitumen and petroleum^{8-11,22}. The predominance of parent PAHs in the boundary samples (Fig. 2) indicates that they have been subjected to far higher temperatures than those involved in petroleum formation. On the other hand, the two samples above and below Woodside Creek K/T boundary are enriched in alkyl

homologues and exhibit patterns characteristic of petroleum^{8-11,22}. However, the effects of biodegradation and water washing²⁴ and weathering²⁵ necessitate caution in interpreting the PAH data. Data from other lipid components indicate that both the Stevens Klint and Woodside Creek samples are biodegraded, although not to the extent of total removal of *n*-alkanes (J. D. *et al.*, preprint). Weathering could cause the removal of alkylated PAHs²⁵, resulting in a pyrolytic profile. However, comparison of the unresolved complex mixture of aliphatic-fraction and normal-alkane profiles from Woodside Creek indicates that the three samples (K/T, +250 cm, -275 cm) apparently were subjected to similar extents of weathering (J. D. *et al.*, preprint; M. I. V., unpublished data). Yet samples above and below the boundary exhibit PAH alkyl homology significantly different from that at the boundary, suggesting that the boundary experienced an input of combustion PAHs significantly above the background level. It is also known that parent PAHs are more susceptible to biodegradation than are the alkyl PAHs and that the rate of biodegradation decreases with increasing degree of alkylation under oxic conditions²⁴. Assuming that some biodegradation has occurred, the elevated abundance of parent PAHs relative to alkyl PAHs at the boundary provides convincing evidence for an input of combustion material to the boundary layer.

Alkyl homology at the boundary is similar to that in the profiles reported from soot¹¹ and recent sediments receiving airborne combustion inputs⁸⁻¹¹ (Figs 2 and 3). The alkylation

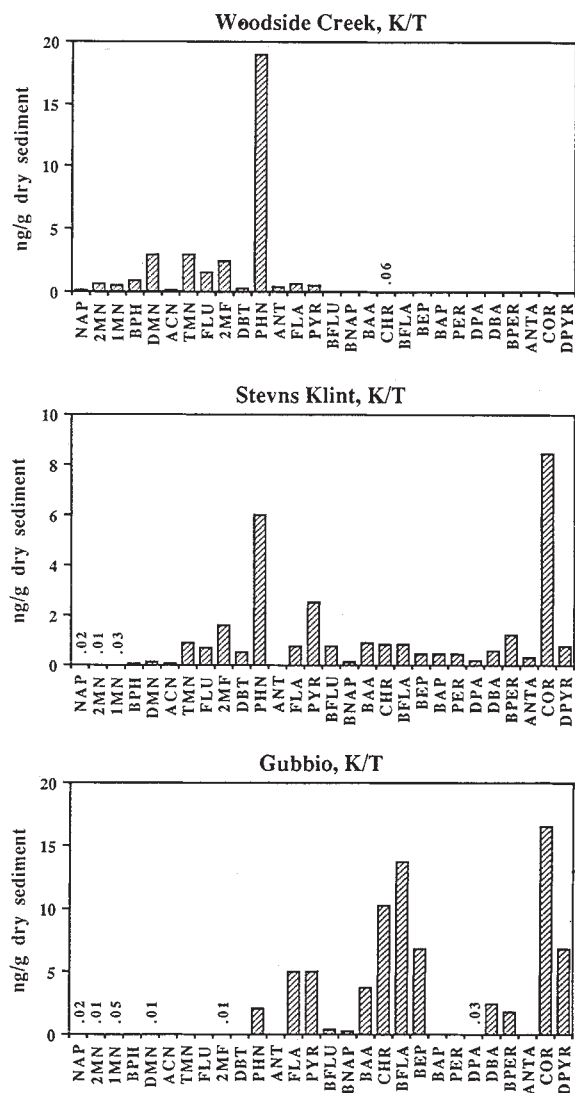


FIG. 1 PAH distribution at the K/T boundaries. The following compounds in the external standard mixture were quantified: naphthalene (NAP); 2-methylnaphthalene (2MN); 1-methylnaphthalene (1MN); biphenyl (BPH); 2,6-dimethylnaphthalene (DMN); acenaphthene (ACN); 2,3,5-trimethylnaphthalene (TMN); fluorene (FLU); 2-methylfluorene (2MF); dibenzothiophene (DBT); phenanthrene (PHN); anthracene (ANT); fluoranthene (FLA); pyrene (PYR); 2,3-benzofluorene (BFLU); 1,1'-binaphthalene (BNAP); benz(a)anthracene (BAA); chrysene/triphenylene (CHR); benzo(b+k)fluoranthene (BFLA); benzo(e)pyrene (BEP); benzo(a)pyrene (BAP); perylene (PER); 9,10-diphenylanthracene (DPA); 1,2,5,6-dibenzanthracene (DBA); benzo(g,h,i)perylene (BPER); anthanthrene (ANTA); coronene (COR); 1,2,4,5-dibenzopyrene (DPYR). Numbers printed above some compounds refer to concentrations too low to be plotted. Quantification was carried out with a Finnigan Model 4000 quadrupole mass spectrometer interfaced directly with a Finnigan Model 9610 gas chromatograph. A fused-silica DB5 (J&W) column of 30 m $\times$ 0.25 mm was programmed from 35 $^{\circ}$ C to 290 $^{\circ}$ C at 4 $^{\circ}$ C min⁻¹ and then held at constant temperature for ~2 hours. Helium was used as a carrier gas; electron energy was 70 eV; source temperature was 240 $^{\circ}$ C; scan speed 2 s per scan from 50 to 550 AMU. Analyses were performed in the full scan mode and the extracted-ion-current profiles were used in the identification and quantification. The PAH fraction was also screened for substitution by alkyl homologues up to C₄ to C₅ (from NAP to PER), in addition to those listed above.

patterns of K/T samples from Woodside Creek and Gubbio are characteristic of combustion products of wood or kerosene, whereas the Stevns Klint profile is more akin to that from combustion of coal²². These patterns can originate from combustion at temperatures of 400–800 °C (as occur during forest fires), which could allow some alkylated PAHs to survive while thermally cracking extended chains into radicals, which recombine to form larger aromatic systems by pyrosynthesis⁸.

Based on the detection of retene (1-methyl-7-isopropyl-

phenanthrene) in Woodside Creek⁵ and in the DSDP site 605 (ref. 26), biomass is considered a strong candidate as the fuel for such a fire⁵. We detected retene and simonellite at <1 part per 10^9 in the K/T samples, indicating that although biomass may have been a fuel source, the retene concentration is too low to confirm this. Alkylated phenanthrenes have indeed been detected in volcanic ash from Mt St Helens as a result of pyrolysis of plant material²⁷, and in emissions from wood combustion²⁰. The yield of retene apparently depends on oxygen availability

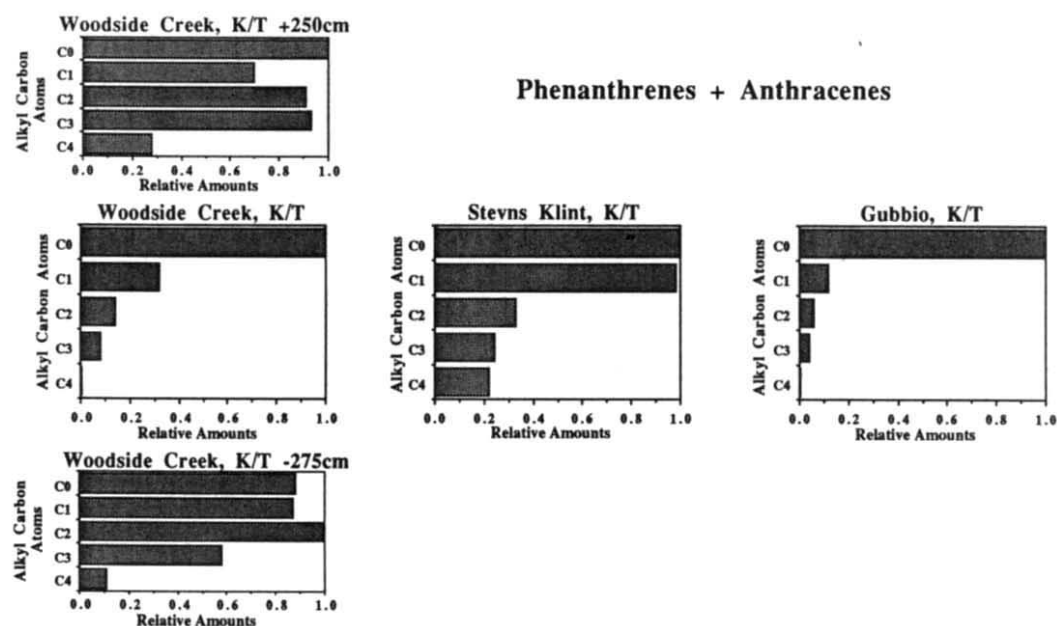


FIG. 2 Alkyl homologue distributions of selected PAHs. The most abundant homologue in each sample is set to 1.0.

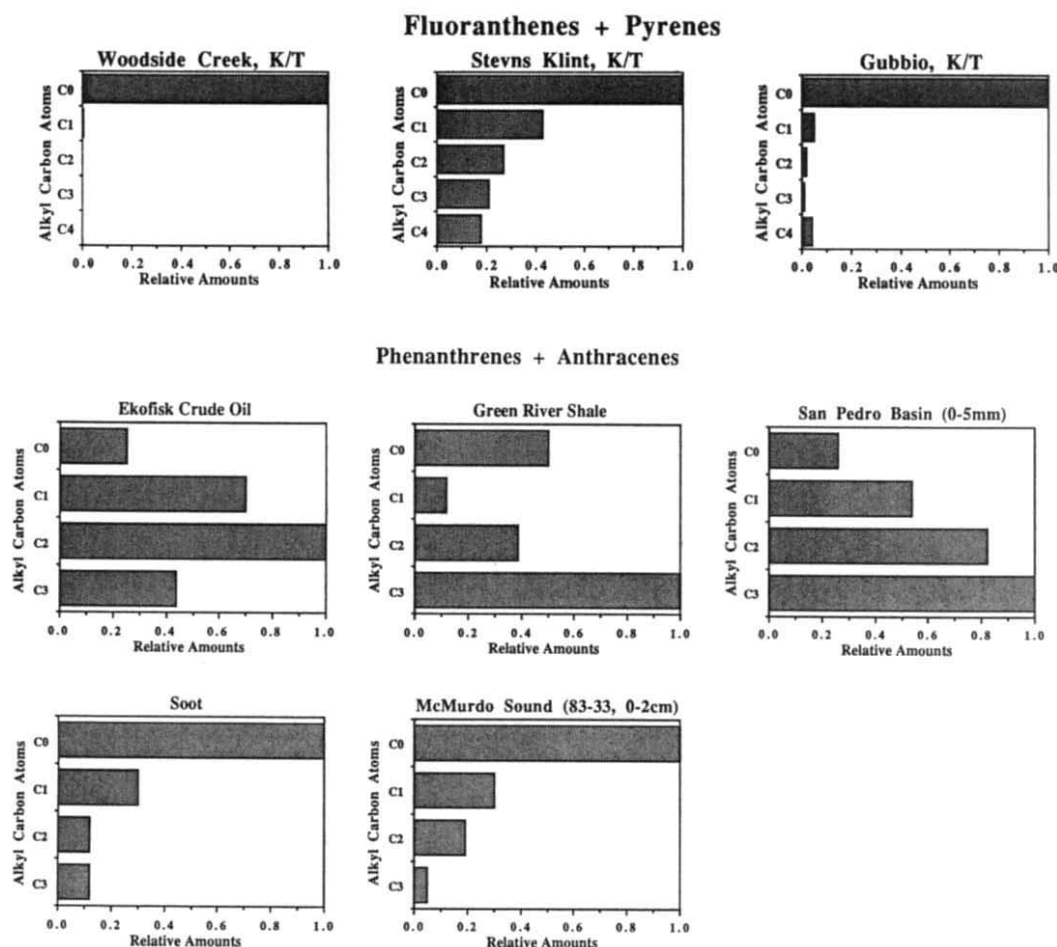


FIG. 3 Alkyl homologue distributions of phenanthrenes + anthracenes plotted from selected literature data. Note the enrichment of alkylated PAHs in oil⁴¹, shale and recent, petroleum-contaminated sediment from San Pedro Basin, southern California Bight⁴⁰. Recent sediment from McMurdo Sound²⁹, Antarctica, illustrates the combustion PAH profile, similar to that found in soot. Airborne combustion characteristics of PAHs in the Antarctic sediments are described in ref. 29. Soot data are plotted from ref. 11. Note the similarity of K/T boundary samples to McMurdo Sound PAH profile, and those of the Woodside Creek K/T +250 and -275-cm to the San Pedro Basin PAH profile.

and is inversely proportional to temperature²⁰. The relatively low ratio of retene to other PAHs (such as phenanthrene) in the K/T boundary clays is consistent with moderate- to high-temperature (400–800 °C) pyrosynthesis⁶. However, the low amounts of retene could also reflect a contribution from diagenesis of dehydroabietic acid²⁸, which would bring into question a thermal origin of retene from biomass fires.

In conclusion, our data suggest that the PAH distributions at the K/T boundary from Woodside Creek, Stevns Klint and Gubbio are characteristic of a combustion origin, which is consistent with the earlier suggestions of massive global fires^{4,5}.

Received 3 November 1988; accepted 27 January 1989.

1. Alvarez, W., Alvarez, L. W., Asaro, F. & Michel, H. V. *Science* **223**, 1183–1186 (1984).
2. Officer, C. B. & Drake, C. L. *Science* **227**, 1161–1167 (1985).
3. Hallam, A. *Science* **238**, 1237–1242 (1987).
4. Wolbach, W. S., Lewis, R. S. & Anders, E. *Science* **230**, 167–170 (1985).
5. Wolbach, W. S., Gilmour, I., Anders, E., Orth, C. J. & Brooks, R. R. *Nature* **334**, 665–669 (1988).
6. Neff, J. M. *Polycyclic Aromatic Hydrocarbons in the Aquatic Environment: Sources Fates and Biological Effects* (Applied Science, London, 1979).
7. Commins, B. T. *Atmos. Environ.* **3**, 565–572 (1969).
8. Youngblood, W. W. & Blumer, M. *Geochim. cosmochim. Acta* **39**, 1303–1314 (1975).
9. Hites, R. A., Laflamme, R. E. & Windsor, J. G., Jr, in *Advances in Chemistry Series 185* (eds Petrakis, L. & Weiss, F. T.) 289–311 (*Am. chem. Soc.*, Washington, DC, 1980).
10. Venkatesan, M. I., Ruth, E. & Kaplan, I. R. *Am. chem. Soc. 192nd Mtg.* Anaheim, California, September 1986.
11. Sporstøl, S. et al. *Environ. Sci. Technol.* **17**, 282–286 (1983).
12. Schmitz, B. *Geology* (in the press).
13. Strong, C. P. *NZ J. Geol. Geophys.* **20**, 687–696 (1977).
14. Brooks, R. R. et al. *Science* **226**, 539–542 (1984).
15. Crockett, J. H., Officer, C. B., Wezel, F. C. & Johnson, G. D. *Geology* **16**, 77–80 (1988).
16. Venkatesan, M. I., Ruth, E., Steinberg, S. & Kaplan, I. R. *Mar. Chem.* **21**, 267–299 (1987).
17. Schmitz, B., Andersson, P. & Dahl, J. *Geochim. cosmochim. Acta* **52**, 229–236 (1988).
18. Blumer, M. *Chem. Geol.* **16**, 245–256 (1975).
19. Kamens, R. M., Guo, Z., Fulcher, J. N. & Douglas, A. B. *Environ. Sci. Technol.* **22**, 103–108 (1988).
20. Ramdahl, T. *Nature* **306**, 580–582 (1983).
21. Simoneit, B. R. T. & Lonsdale, P. F. *Nature* **295**, 198–202 (1982).
22. Laflamme, R. E. & Hites, R. A. *Geochim. cosmochim. Acta* **42**, 289–303 (1978).
23. Gilmour, I. & Guenther, F. *Conf. on Global Catastrophes in Earth History* Snowbird, Utah, 60–61 (1988).
24. Volkman, J. K., Alexander, R., Kagi, R. I., Rowland, S. J. & Sheppard, P. N. *Org. Geochem.* **6**, 619–632 (1984).
25. Clayton, J. L. & King, J. D. *Geochim. cosmochim. Acta* **51**, 2153–2157 (1987).
26. Simoneit, B. R. T. & Beller, H. R. *Init. Rep. DSDP 93*, Ch. 52 (eds van Hinte, J. E. et al.) 1211–1215 (US Govt Printing Office, Washington, DC, 1987).
27. Pereira, W. E., Rostad, C. E., Taylor, H. E. & Klein, J. M. *Environ. Sci. Technol.* **16**, 387–396 (1982).
28. Simoneit, B. R. T. *Geochim. cosmochim. Acta* **41**, 463–476 (1977).
29. Venkatesan, M. I. *Org. Geochem.* **12**, 13–27 (1988).

ACKNOWLEDGEMENTS. We thank Professors R. R. Brooks, W. Cavazza and B. Schmitz for the samples, E. Ruth for GC/MS analyses and F. Kytte for reviewing the manuscript. We also appreciate reviews by I. Gilmour and J. Smit. This work was partially supported by the OHER division of DOE and UERG's Energy, Science and Technology Program.

Graphitized diamonds from a peridotite massif in Morocco and implications for anomalous diamond occurrences

D. G. Pearson*, G. R. Davies*, P. H. Nixon* & H. J. Milledge†

* Department of Earth Sciences, University of Leeds, Leeds LS2 9JT, UK

† Department of Geological Sciences, University College London, Gower Street, London WC1E 6BT, UK

DIAMONDS are found in several areas of the world where their currently accepted source, kimberlite/lamproite volcanism, has not been found. Garnet pyroxenite layers in the Beni Bousera peridotite massif, northern Morocco, contain octahedral and other cubic forms of graphite, which we interpret here as pseudomorphs after diamond. This occurrence demonstrates that fragments of highly diamondiferous mantle are tectonically emplaced into the continental crust, providing an alternative source for some diamonds.

The Beni Bousera peridotite massif forms part of the Betico-Rifean Fold Belt. A continuous gravity high between Beni

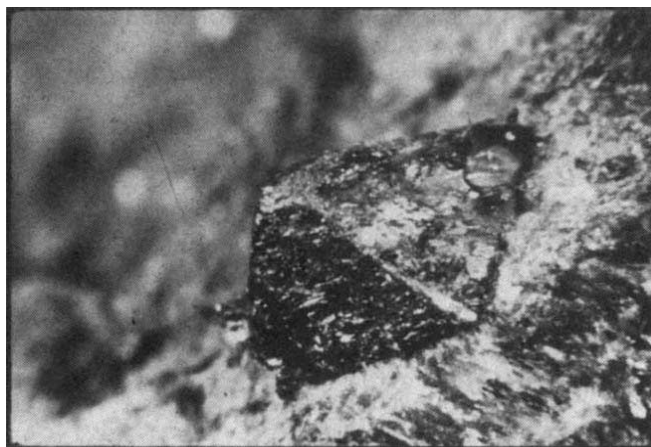


FIG. 1 Photograph of upper half of a sharp-edged octahedron protruding from the garnet pyroxenite layer.

Bousera and the Ronda peridotite in south-west Spain¹ indicates that the Betics and Rif are part of the same tectonic system. Both the Beni Bousera and Ronda massifs were emplaced into the crust at high temperature during the Alpine orogeny^{2,3}. In addition to peridotites, both massifs contain many mafic layers of varying composition. The predominant rock type within the Beni Bousera peridotite massif is a variably altered spinel lherzolite accompanied by rare harzburgites and dunite pods. Mafic layers form up to 10% of the outcrop⁴. Four garnet clinopyroxenite layers are known to contain graphite octahedra⁵ (Fig. 1). The graphite pyroxenite layers comprise orange pyrope almandine garnet and omphacitic pyroxene porphyroclasts with minor amounts of plagioclase, spinel and sulphides. Graphite forms up to 15% by volume of the layers, and never cross-cuts the layer margins into the surrounding peridotites, which significantly contain no disseminated graphite. Graphite is also found in late-stage, low-temperature Cu–Ni mineralized veins which cross-cut the emplacement fabric of the massif. The graphite-garnet clinopyroxenites are not associated with this mineralization.

Graphite forms liberated by acid dissolution are multicrystalline assemblages of small graphite crystallites, here termed aggregates. They consist of (1) sharp-edged octahedra, with or without rounded fibrous graphite coats (the most abundant form); (2) rhombicuboctahedra⁶ exhibiting well formed cube and octahedral faces; (3) contact twins or macles, with and without re-entrant angles, most of which are coated; (4) irregular to rounded masses whose surface morphology resembles framesite (phanerocrystalline carbonado⁷); and (5) octahedra that have been flattened and/or sheared along the {111} faces.

Graphite with no regular symmetry or form is widespread and probably represents highly deformed cubic forms. The regular crystalline forms of this graphite belong to the cubic system whereas natural graphite crystallizes in the hexagonal system. Graphite could be pseudomorphic after several cubic minerals. The most plausible geologically are diamond and spinel group minerals such as chromite and magnetite. Sub-millimetre anhedral spinel grains do occur in some mafic layers in Beni Bousera (less than 1% by volume). However, the relatively low Cr content (450–1,000 p.p.m.) of the bulk rock, and the low Cr and Fe³⁺ content of the silicate minerals, are not consistent with the rock having once contained up to 15% of either chromite or magnetite. Additionally, petrographic studies show no evidence of either mineral as relicts within the graphite aggregates, and ashed graphites⁸ confirm their absence.

All the regular crystalline forms of the graphite are displayed by natural diamond. Scanning electron microscopy (SEM) reveals that {111} faces of the graphite octahedral cores display pronounced steps which are comparable to the {111} growth

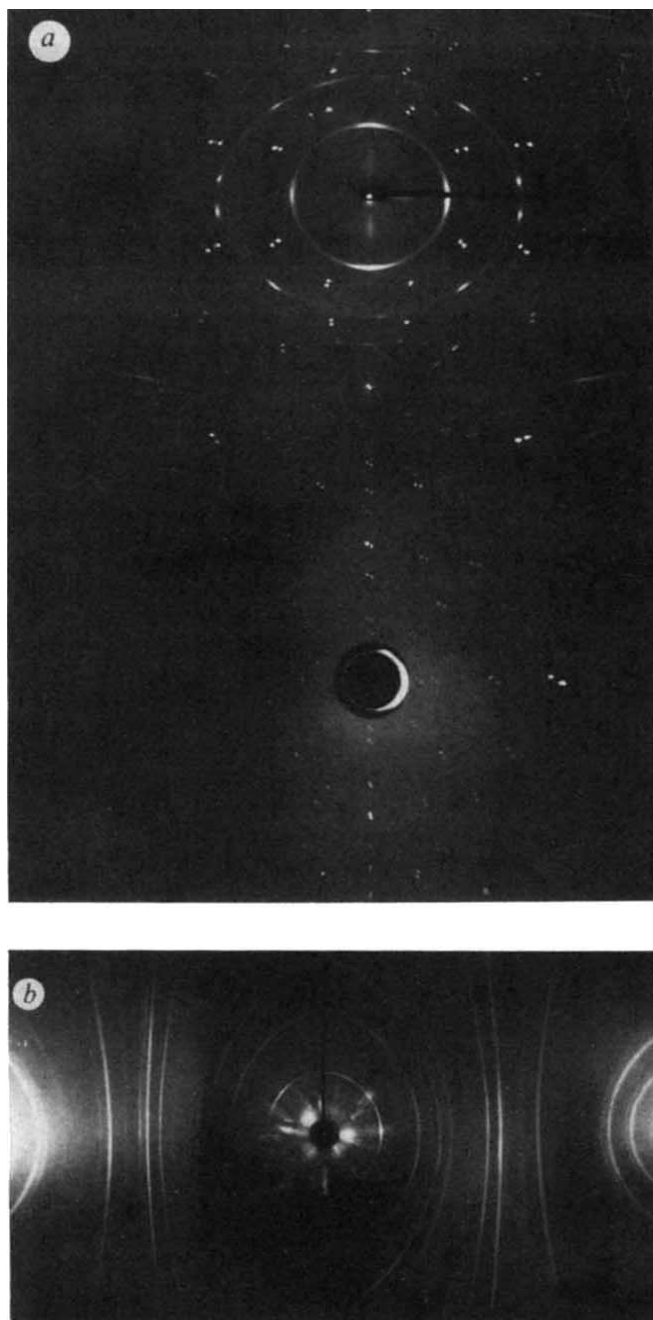


FIG. 2 *a*, Laue photograph of experimentally graphitized diamond, showing the strongly preferred orientation of graphite [0001] parallel to diamond {111} (courtesy of G. Cooper). Cu $K\alpha$ radiation. [110] axis vertical; beam along [001]. *b*, Laue photograph of Beni Bousera graphite octahedron, showing preferred orientation of graphite [0001] parallel to {111} of the octahedron. Cu $K\alpha$ radiation. The [110] axis is vertical; beam positioned to show the expected enhancement of graphite (0002) on the zero layer assuming it is parallel to {111}.

plates observed on natural diamond. Also evident on the {111} faces are small pits of trigonal symmetry, which have similar geometry and orientation to the etched trigons on the faces of diamond from volcanic pipes. In addition, SEM observations demonstrate that the basal {0001} planes of graphite crystals in the core are approximately parallel to the {111} faces of the octahedra. In contrast, graphite crystals in the fibrous coat are much more irregular in their orientation, with their basal planes at a high angle, often perpendicular, to the {111} surfaces of the core. X-ray diffraction^{8,9} (G. Cooper, personal communica-

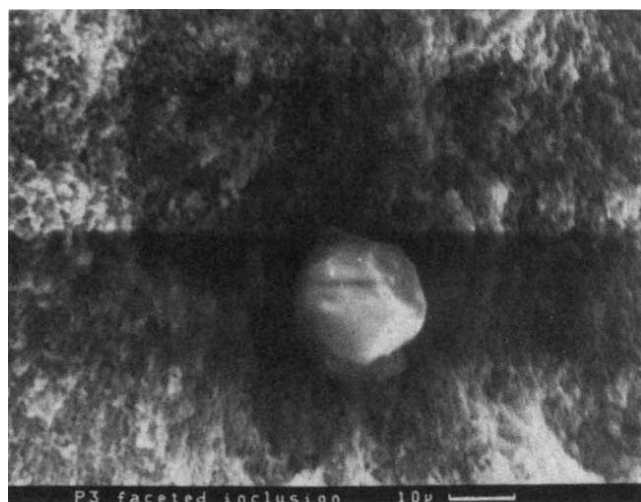


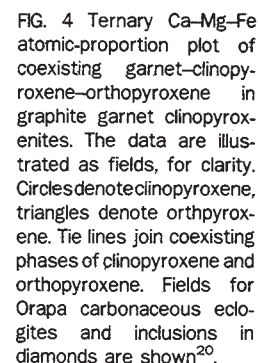
FIG. 3 Scanning electron photomicrograph of a clinopyroxene inclusion within a graphite octahedron, displaying apparent cubo-octahedral faceting. The {100} and {111} facets of the inclusion are parallel to the {100} and {111} faces of the octahedron.

tion) reveals that the graphite produced by experimental graphitization of natural diamond between 4.8 and 10 GPa pressure produces graphite crystallites whose {0001} planes lie parallel to the {111} faces of the original diamond octahedra (Fig. 2*a*). Single-crystal X-ray diffraction patterns of the octahedral graphite forms from Beni Bousera sometimes show a strong preferred orientation similar to that observed on experimentally graphitized diamond (Fig. 2*b*).

Graphite exhibiting cubic morphology has been found previously only in extraterrestrial samples, in the form of cliftonite¹². In these cases the multicrystalline graphite aggregates take the form of cubes, cubo-octahedra and pseudo-hexagonal outlines. These aggregates do not display the preferred orientation noted above. The sixfold axis of the graphite lies parallel to the [001] and [113] axes of the cliftonite^{13,14} and thus these examples are not considered to be pseudomorphs after diamond^{13,14}. Therefore, graphite pseudomorphing a cubic mineral will not necessarily acquire the preferred orientation seen in the Beni Bousera graphite.

Clinopyroxene and garnet are both observed as inclusions within octahedra in thin section. Radiographs indicate that these inclusions in the graphite octahedra occasionally occupy a large proportion of the interior. Some inclusions possess well formed cubo-octahedral facets identical to many inclusions within natural diamond^{10,11} (Fig. 3). The mechanism by which diamond imposes cubic symmetry on minerals of lower symmetry, such as monoclinic pyroxene, is not clearly understood. However, this phenomenon is powerful evidence for a diamond precursor for the Beni Bousera graphite aggregates.

The Beni Bousera octahedral graphite aggregates display a sharp physical boundary between cores of cubic symmetry with regular crystallites, and rounded fibrous coats. In general, there is a difference in style of deformation between core and coat. Core graphite exhibits relatively coarse sub-grains when cut parallel to a cube axis, whereas the fibrous-coat graphite possess kink lamellae. These differences are due to the marked mechanical anisotropy of graphite, deformation style depending on the orientation of graphite basal planes with respect to the direction of principal stress. Raman spectroscopy (J. D. Pasteris, unpublished data) indicates that both coat and core graphite are extremely well crystalline and hence of high-temperature origin. The graphite aggregates are thus similar to coated diamonds; the latter are characterized by well formed cores, often clear octahedra, surrounded by diamond displaying a fibrous, radial growth, giving rise to the generally rounded



All of these lithologies contain evidence of lower-pressure re-equilibration. Garnet, orthopyroxene and plagioclase exsolution occur within clinopyroxene. There is no major- or trace-element compositional difference between garnet and clinopyroxenes included in graphite octahedra and those form-

Obducted ophiolites in Tibet²² and Alpine peridotites in the Koryuk Mountains of the eastern Soviet Union and Armenia²³ are reported to contain very low concentrations of diamond. Alluvial diamonds, as yet of unexplained provenance but situated in the region of ophiolites or similar ultramafic belts including the Urals²³, the Appalachians, Kalimantan (Borneo) and Copeton-Bingara (New South Wales), have no known volcanic sources²⁴. The source rocks of these 'anomalous diamond occurrences', whether proven or suspected, are not eclogites but are mainly highly magnesium chromite harzburgites. These are similar to diamondiferous peridotite xenoliths in kimberlites except that garnet is usually lacking. Thus, rocks comparable to diamondiferous eclogite and peridotite xenoliths in kimberlite occur in some of the Earth's major collision zones, having been tectonically emplaced from close to the base of the lithosphere. Collision zones containing ultramafic massifs thus provide a potential economic source of diamonds. □

1. Bonini, W. E., Loomis, T. P. & Robertson, J. D. *J. geophys. Res.* **78**, 1372-1382 (1973).
2. Dickey, J. S. *Miner. Soc. Am. spec. Pap.* **33**, 43-49 (1970).
3. Loomis, T. P. *Am. J. Sci.* **275**, 1-30 (1975).
4. Kornprobst, J. *Contr. Miner. Petrol.* **23**, 283-322 (1969).
5. Slodkevich, V. V. *Zapiski Vses. Miner. Ob-Va.* 13-33 (in Russian) (1980).
6. Moore, M. *Ind. Diamond Rev.* **2** (1985).
7. Jaynes, C. *Ind. Diamond Rev.* 14-23 (1978).
8. Lonsdale, K. & Milledge, H. J. in *The Physical Properties of Diamond* (ed. Berman, R.) (Oxford University Press, 1965).
9. Evans, T. in *Properties of Diamond* (ed. Field, J. E.) 425-469 (Academic, London, 1979).
10. Mitchell, R. S. & Giardini, A. A. *Am. Miner.* **38**, 136-138 (1953).
11. Mendelssohn, M. J., Milledge, H. J., Nave, E. & Woods, P. A. *Acta crystallogr.* **A31**, S214 (1975).
12. Haidinger, W. K. & Partsch, P. *Ann. Phys.* **67**, 437-439 (1846).
13. Grenville-Wells, H. J. *Miner. Mag.* **29**, 803-816 (1952).
14. Okada, A. & Shima, M. *J. Jap. Ass. Miner. Petrol. Econ. Geol.* **67**, 45-49 (1972).
15. Custers, J. F. H. *Am. Miner.* **35**, 51-58 (1955).
16. Kamiya, Y. & Lang, A. R. *Phil. Mag.* **11**, 347-356 (1965).
17. Robinson, D. N. *Miner. Sci. Engng* **10**, 55-72 (1978).
18. Gurney, J. J. *Proc. 4th Int. Kimberlite Conf.* (in the press).
19. Robinson, D. N. thesis, Univ. Capetown (1979).
20. Robinson, D. N., Gurney, J. J. & Shee, S. R. in *Developments in Petrology* Vol. 11B, (ed. Kornprobst, J.) 11-24 (Elsevier, Amsterdam, 1984).
21. Moore, R. O. & Gurney, J. J. *Nature* **318**, 553-555 (1985).
22. Fang, Chingsong & Bai Wenji *Int. geol. Rev.* **27**, 455-457 (1981).
23. Kaminskii, F. V. *Zapiski Uses. Miner. Ob.* **4**, 488-493 (1980).
24. Nixon, P. H., Davies, G. R., Slodkevich, V. V. & Bergman, S. C. *Extended Abstracts, 4th Int. Kimberlite Conf.* 412 (1986).

ACKNOWLEDGEMENTS. We have benefitted from discussions with Drs S. R. Boyd, J. J. Gurney, J. W. Harris, D. P. Mathey and V. V. Siodkevich. The co-operation of the Ministère de l'Énergie et des Mines, Rabat, is gratefully acknowledged. We thank Drs A. Tindle and P. Hill for probe analyses and Drs N. Shimizu and J. Pasteris for providing unpublished data. D.G.P. and G.R.D. gratefully acknowledge NERC funding.

Coagulation on bubbles allows microbial respiration of oceanic dissolved organic carbon

Paul E. Kepkay & Bruce D. Johnson*

Biological Oceanography Division, Department of Fisheries and Oceans, Bedford Institute of Oceanography, Dartmouth, Nova Scotia B2Y 4A2, Canada

* Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, B3H 4J1, Canada

DISSOLVED organic carbon in sea water (DOC) is one of the chief reservoirs of reactive organic carbon on the planet¹. To determine the rate at which this carbon breaks down, a long-standing paradox must be solved. DOC appears to be remarkably unreactive², yet it must be reactive to maintain the mass balance between organic carbon in the ocean and CO₂ in the atmosphere¹. Here we show that the coagulation of colloidal DOC on bubble surfaces initiates the rapid microbial respiration of carbon which would otherwise be less accessible to the biota. This coupling of respiration to surface coagulation as a physical means of regenerating a substantial fraction (5–15%) of oceanic DOC could be a key factor in the mechanism required to recycle a recalcitrant reservoir of carbon back to CO₂.

During August and September 1988, sea water was collected from a depth of 10 m on a transect of four stations across a highly productive tidal front on Georges Bank, and from four depths (from 10–85 m) at a station occupied for 9 days in the Sargasso Sea. On board ship all the sea water samples were filtered through 2- μ m Nuclepore membranes to remove the majority of particles without removing colloidal organic material^{3,4,5} and were bubbled for 30 min in a glass column, using a 10–25 μ m glass frit to produce bubbles with an average diameter of 487 μ m (ref. 5). The air supplied to the frit was prefiltered through activated carbon and a sterile 0.2- μ m filter cartridge (Gelman) to ensure that contaminants were not introduced on bubble surfaces. After 30 min of bubbling, the sea water was divided into duplicate two-litre batches which were gently stirred in an incubator set at the appropriate *in situ* temperature. Two additional two-litre batches of filtered sea water were stirred at the same *in situ* temperature to act as unbubbled controls. Samples from all batches were taken at intervals to determine the rate of oxygen consumption by an oxygen gradient technique⁶.

The concept that aggregates of organic material and bacteria are generated by the coagulation of DOC on bubble surfaces has been discussed for over 25 years^{5,7}. In our experiments this process of surface coagulation caused a rapid increase in microbial respiration, as measured by oxygen consumption rate (Figs 1 and 2). Respiration reached a maximum after 2–4 h and then declined to values similar to those found in unbubbled controls by 8–20 h. This is the first demonstration that surface coagulation can have a large, reproducible and systematic effect on microbial respiration in the open ocean—not only from the highly productive Georges Bank⁸, but also from a far less productive Sargasso Sea⁹. Without rapid sampling at the beginning of the experiments, we would have missed the dynamics associated with a transient microbial community and inferred that bubbling had no effect.

Many bubbling experiments have failed to show that surface coagulation has a significant effect on microbial communities⁵. Our experiments succeeded because we paid close attention to results obtained by Batoosingh *et al.*³, Johnson and Wangersky⁴, and Johnson *et al.*⁵. Filtering sea water through anything finer than a 2- μ m pore size removes larger colloids in the 0.2–1.2 μ m size range—a primary source of labile organic material for

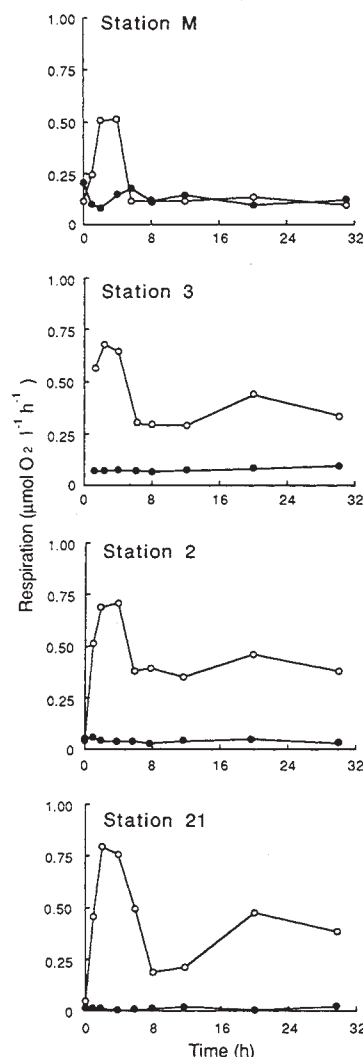


FIG. 1 Respiration in bubbled sea waters (open circles) and unbubbled controls (closed circles) from a depth of 10 m at four stations on the NE quadrant of Georges Bank (47° 00'N; 67° 51'W). The transect of 4 stations included surface-water environments ranging from the mixed layer of a stratified water column in the Fundian Channel (station 21) to a tidal front at the edge of the bank (stations 2 and 3) and finally to well-mixed, particle-rich water on the Bank itself (station M). Respiration was measured using an oxygen gradient technique⁶. Samples (10 ml) were pipetted into polycarbonate filter manifolds and filtered through 2- μ m Nuclepore filters. The same volume of filter-sterilized (0.2- μ m filtered) sea water was then added to each manifold and oxygen gradients were measured at *in situ* temperature to a distance of 5 mm above the material collected on the surface of each filter. The gradients allowed calculation of the downward flux of oxygen over a given length of time and thus the rate of consumption per cm² of surface or per ml of original sample. Error bars from selected triplicate incubations were all smaller than the dimensions of the symbols.

surface coagulation⁵. Physical models of the mass transfer of organic particles in turbulence¹⁰ show that colloids of about 1 μ m in diameter are situated at the boundary between two domains of mass transfer. Particles smaller than 1 μ m are more rapidly transported to bacterial cells by convective diffusion, whereas larger particles are more effectively transported by collision with cells in turbulent shear. In effect, there appears to be an untapped reservoir of labile organic carbon that is not easily accessed by bacteria in the absence of surface coagulation. This in turn leads to an obvious but important conclusion. The interpretation of data from experiments designed to illustrate the microbial effects of bubbling will be seriously compromised if colloidal organic material has been removed by filtering.

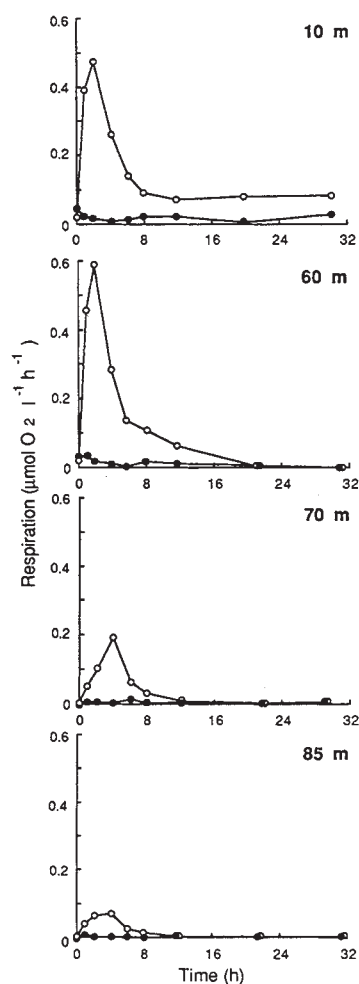


FIG. 2 Respiration in bubbled sea waters (open circles) and unbubbled controls (closed circles) from four depths at a station occupied for 9 days in the Sargasso Sea (at 36° 00'N; 64° 05'W). During this time, the mixed layer of surface water extended down to a depth of 60–70 m. Error bars from selected triplicate incubations were all smaller than the dimensions of the symbols.

On Georges Bank, the mean respiration in unbubbled controls increased (Fig. 1; Table 1) as the environment at a depth of 10 m progressed from the mixed layer of a stratified water column in the Fundian Channel (station 21) to a tidal front at the edge of the Bank (stations 2 and 3) and finally to well mixed, particle-rich water on the Bank itself (station M). Respiration was enhanced by surface coagulation to approximately the same absolute value at all 4 stations but, when compared with the controls, enhancement decreased by over an order of magnitude from station 21 to station M (Table 1). The reduced effect of surface coagulation in well mixed water at station M (ref. 11) was probably related to increased respiration of labile organics coagulated with larger numbers of suspended particles (1.1 p.p.m. at station M compared with 0.2 p.p.m. at station 21; P.E.K. and D.K. Muschenheim, unpublished observations). As a result, the effects of bubbling and surface coagulation may be of greater biological significance in sea water with low concentrations of suspended particulates, such as that found at station 21.

Bubbling enhanced respiration by a factor of 4.8–11.2 in the Sargasso Sea (Fig. 2; Table 1). This enhancement was positively correlated with primary production (Fig. 3) and decreased rapidly below the base of the mixed layer at 60–70 m (Fig. 2). Surface coagulation in this particle-poor environment may therefore be confined in its biological effect to relatively 'new' colloids associated with primary production in the upper ocean. The reasons why surface coagulation caused such a rapid microbial response are not clear, but the answer may lie in the results of Baylor *et al.*¹² and Graham *et al.*¹³, which show that bubbling sea water for 18–45 min can concentrate phosphorus into the adsorbed or coagulated phase by a factor of 4 to 200. Given that microbial growth is probably not carbon-limited in surface waters¹⁴, the degree to which nutrients such as phosphorus and nitrogen are involved in surface coagulation may ultimately regulate the magnitude of a microbial response to aggregate generation.

It still remains to be seen if bubble spectra in a glass column are typical of those found under a breaking wave, but it is our experience that data from bubbling columns may underestimate, rather than overestimate, the physical and biological effects of surface coagulation. A bubbling column can appreciably enhance the transport of 1 μm colloids to bacteria, but models of mass transfer of the same material by natural bubble populations under breaking waves¹⁰ indicate that the enhancement may

TABLE 1 Mean respiration of bubbled sea waters and unbubbled controls in relation to previous measurements from Georges Bank and the Sargasso Sea

Location	Depth (m)	Respiration (unbubbled sea water)		Respiration (bubbled sea water)		Enhancement factor†
		Mean*	Range	Mean*	Range	
Georges Bank (this study)						
Station M	10	100	62–158 (n=9)	359	146–783 (n=9)	2.7
Station 3	10	58	54–73 (n=9)	369	223–523 (n=9)	6.4
Station 2	10	31	19–46 (n=9)	334	61–546 (n=9)	11.0
Station 21	10	9	3–15 (n=9)	323	68–510 (n=9)	35.8
Georges Bank ²¹	surface	185	6–288 (n=6)	—	—	—
Georges Bank and Gulf of Maine ²²	1–10	132	3–190 (n=10)	—	—	—
Sargasso Sea (this study)	10	12	5–28 (n=9)	134	18–364 (n=9)	11.2
	60	15	5–34 (n=8)	129	3–450 (n=8)	8.6
	70	6	2–12 (n=9)	40	3–148 (n=9)	6.7
	85	4	2–9 (n=8)	19	2–53 (n=9)	4.8
Sargasso Sea ²³	surface	118	5–230 (n=13)	—	—	—
Sargasso Sea ²⁴	20–75	90	40–230 (n=4)	—	—	—
Bedford Basin (this study)	10	31	15–46 (n=8)	361	31–576 (n=8)	11.8

Respiration in the unbubbled controls fell within the range of values from earlier work and respiration in many of the bubbled sea waters exceeded the range of values from earlier work. The systematic variations in respiration that we observed in the presence or absence of bubbling may therefore have already been observed in earlier studies as apparently random variations. See ref. 20 for previous measurements.

* Mean respiration in the bubbling experiments was calculated from 8 or 9 oxygen consumption rates measured over a 32-h period.

† The enhancement of respiration was calculated from mean respiration (bubbled) divided by the mean respiration from an equivalent unbubbled control.

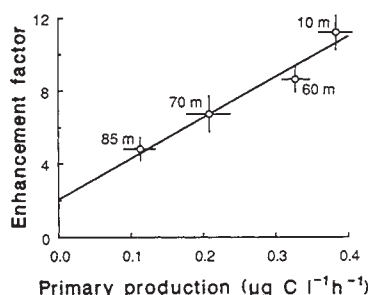


FIG. 3 Linear correlation of the enhancement of mean respiration by surface coagulation (Table 1) and *in situ* primary production at four depths in the Sargasso Sea (W. G. Harrison and B. Irwin, unpublished results). The equation of the line best fitting the data is $y = 2.3 + 23.8x$ ($r = 0.99$).

be too low by a factor of 2–20. This reinforces the concept that a substantial fraction of labile DOC may escape degradation by virtue of its particle size characteristics, and that surface coagulation is required to make the carbon available to bacteria.

The rapid respiratory responses in Figs 1 and 2 leave no doubt that the aggregates produced by coagulation are labile. When cumulative oxygen uptake was calculated from consumption rates on Georges Bank, 10–13 $\mu\text{mol O}_2 \text{ l}^{-1}$ were consumed over 32 h in bubbled sea waters from 10 m. This was equivalent to the respiration of 0.12–0.16 mg C l^{-1} , or 10–15% of the 1.1 mg C l^{-1} of DOC available for coagulation, as estimated from DOC analyses using the ultraviolet oxidation technique of Cauwet¹⁵ (P.E.K., unpublished results). There is some uncertainty^{16,17} and controversy¹⁸ involved in the interpretation of DOC measurements in sea water, but the respiration of 10–15% of a given DOC pool in as little as 32 h indicates that significant quantities of carbon could be made available to the biota by surface coagulation. Even if the measurement controversy is resolved and there is more DOC than was previously thought in the upper ocean¹⁷, there is still a wide variation in the quantity of material coagulated by bubbling. Only 5% of the DOC pool (0.035 mg C l^{-1} from a pool size of 0.7 mg C l^{-1} ; P.E.K., unpublished results) was respired over 32 h when Sargasso sea water from 10 m was bubbled. Although this was 2–4 times lower than the carbon respired at the same depth on Georges Bank, primary production (4.5 $\mu\text{g C l}^{-1} \text{ d}^{-1}$ compared with 50–400 $\mu\text{g C l}^{-1} \text{ d}^{-1}$ on Georges Bank; W. G. Harrison and B. Irwin, unpublished results) was 12–88 times lower. This means that the physical regeneration of organic material by coagulation on bubble surfaces may have a more pronounced effect on microbial communities in regions of low productivity like the Sargasso Sea.

The results obtained when sea water from 10 m in Bedford Basin, Halifax Harbour (Canada) was bubbled during June 1988 (Table 1) demonstrate that respiration in a coastal environment can be enhanced by a factor within the range of values from the open ocean. More importantly, bacterial counts by epifluorescence¹⁹ showed that a rapid increase in bacterial number (to values as high as $16 \times 10^6 \text{ cells ml}^{-1}$) was associated with the respiration induced by surface coagulation. Bacterial growth was then followed by a decline as protozoan numbers increased, and the end result of 32 h of intense microbial activity was a relatively low number of bacteria ($1.6 \times 10^6 \text{ cells ml}^{-1}$). Thus, when different amounts and different fractions of DOC are made available as different sea waters are bubbled, the rapidity of the biological response to surface coagulation is a common factor which cannot be ignored. The coagulation of DOC by bubbling may well be an important oceanographic process in its own right^{5,10} but, when combined with rapid respiration of the coagulated material, it could also be a mechanism for recycling a globally important reservoir of carbon¹ back to CO_2 in the upper ocean. □

Received 10 November 1988; accepted 23 January 1989.

- Hedges, J. I. *Nature* **330**, 205–206 (1987).
- Williams, P. M. & Druffel, E. R. M. *Nature* **329**, 246–248 (1987).
- Batoosingh, E., Riley, G. A. & Keshwar, B. *Deep Sea Res.* **16**, 213–219 (1969).
- Johnson, B. D. & Wangersky, P. J. *Limnol. Oceanogr.* **30**, 966–971 (1985).
- Johnson, B. D., Zhou, X. L. & Wangersky, P. J. *Neth. J. Sea Res.* **20**, 201–210 (1986).
- Kepkay, P. E., Schwinghamer, P., Willar, T. & Bowen, A. J. *Appl. envir. Microbiol.* **51**, 163–170 (1986).
- Blanchard, D. C. in *Air-sea Exchange of Gases and Particles* (Reidel, Boston, 1983).
- O'Reilly, J. E., Evans-Zetlin, C. & Busch, D. A. in *Georges Bank* (MIT, Boston, 1987).
- Menzel, D. W. & Ryther, J. H. *Deep Sea Res.* **7**, 282–288 (1961).
- Johnson, B. D. & Kepkay, P. E. *Limnol. Oceanogr.* (submitted).
- Loder, J. W. & Greenberg, D. A. *Cont. Shelf Res.* **6**, 397–414 (1984).
- Baylor, E. R., Sutcliffe, W. H. & Hirschfeld, D. S. *Deep Sea Res.* **9**, 120–124 (1962).
- Graham, W. F., Piotrowicz, S. R. & Duce, R. A. *Mar. Chem.* **7**, 325–342 (1979).
- Van Es, F. B. & Meyer-Reil, L. A. in *Advances in Microbial Ecology* Vol. 6 (Plenum, New York, 1982).
- Cauwet, G. *Mar. Chem.* **14**, 297–306 (1984).
- Gershey, R. M., MacKinnon, M. D. & Williams, P. J. LeB. *Mar. Chem.* **7**, 289–306 (1979).
- Sugimura, Y. & Suzuki, Y. *Mar. Chem.* **24**, 105–131 (1988).
- Williams, P. M. & Druffel, E. R. M. *Oceanography* **1**, 14–17 (1988).
- Schwinghamer, P. & Kepkay, P. E. *Biol. Oceanogr.* **4**, 289–322 (1987).
- Williams, P. J. LeB. in *Heterotrophic Activity in the Sea IV* (Plenum, New York, 1984).
- Riley, G. A. *Bull. Bingam Oceanogr. Collect.* **7**, 1–74 (1941).
- Packard, T. T. & Williams, P. J. LeB. *Oceanologica Acta* **4**, 351–358 (1981).
- Riley, G. A. *J. Mar. Res.* **2**, 145–162 (1939).
- Williams, P. J. LeB. & Jenkinson, N. W. *Limnol. Oceanogr.* **27**, 576–582 (1982).

ACKNOWLEDGEMENTS. We thank Trevor Platt, Glen Harrison and Bill Li for comments.

Developmental failure and loss of reproductive capacity in the rare palaeoendemic shrub *Dedeckera eurekensis*

Delbert Wiens*, Daniel L. Nickrent†, Cedric I. Davern*‡, Clyde L. Calvin§ & Nancy J. Vivrette||

* Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA

† Department of Plant Biology, University of Illinois, Urbana, Illinois 61801, USA

‡ Department of Biology, Portland State University, Portland, Oregon 97207, USA

|| Ransom Seed Laboratory, PO Box 300, Carpinteria, California 93013, USA

EVOLUTIONARILY successful outcrossing plants have spontaneous seed sets that are, on average, around 50% ($\pm 20\%$), but those of woody species are lower (33%) and herbaceous perennials higher (57%); self-pollinating (autogamous) species average about 85%^{1–3}. *Dedeckera*, an outcrossing (protandrous) evolutionary relict from the Mojave desert of California has a seed set of only 2.5% and no means of vegetative reproduction. We show here that the most probable explanation for this exceptional result is that there are genetically mediated embryonic abnormalities. The percentage of fully viable seeds is further reduced by low viability, germinability, and post-germination developmental failure. Genetic studies involving extinction have emphasized inbreeding depression and homozygosity^{4–8}, but *Dedeckera* is highly heterozygous (27%). Outcrossing does not significantly increase seed set over selfing, suggesting that a high segregational genetic load is primarily responsible for the low seed set in *Dedeckera*. Such loss of reproductive capacity could explain the relictual nature of *Dedeckera* and might ultimately result in its extinction.

A histological analysis of post-fertilization development of 27 ovules of *Dedeckera* showed a pattern of genetically controlled embryo abortion similar to that reported in *Epilobium*³ (Onagraceae), and *Arabidopsis*^{9,10} (Brassicaceae), and in species hybrids¹¹. Three ovules showed no evidence of fertilization.

Spontaneous seed germination in seven trials was only 3.5% ($N = 460$ filled seeds). Sixty days of stratification had no effect, but scarification and gibberellic acid treatment increased germination to 33%. Only 11.1% of spontaneously germinating seeds

‡ Deceased.

developed normally, but the sample size was small ($N=9$). Thus the percentage of fully viable seeds in *Dedeckera* is probably less than 0.5%. *Dedeckera* may produce about 64,500 flowers and ovules (ovaries are uniovulate). Given seedling survivorship values for desert perennials^{12,13} of 10^{-4} to 10^{-5} , the loss of reproductive capacity in *Dedeckera* may be critical for its continued survival.

Neither prezygotic genetic self-incompatibility nor inadequate pollination explain the exceptionally low fecundity of *Dedeckera* because about 90% of the ovaries initiate growth, which is taken as *prima facie* evidence for fertilization (Table 1). The consistency of the data, both between populations and in different years, indicates that resources do not limit seed set (Table 1), and also argues against gametic imbalance and hybrid dysgenesis as explanations for low seed set. *Dedeckera* is not a chromosome translocation heterozygote (n is about 12 bivalents). Herbivory was uncommon and there was no overt evidence of pathogen activity.

Intraplant self-pollination (geitonogamy) may be common in *Dedeckera*¹⁴, but inbreeding depression resulting from selfing in a normally outcrossed population (segregational genetic load) is not a major factor contributing to the low seed set (Table 2). The seed set for controlled self-pollinations (6.3%) and cross-pollinations (11.4%) do not differ significantly ($d.f.=1$, $\chi^2=2.95$, $P>0.05$) and all population sizes exceed 150 (Table 2).

Allozyme analyses were completed from each of the six populations studied (Table 1) to determine the amount of genetic polymorphism and heterozygosity. F -statistics¹⁵ were used to examine apportionment of variation within and between populations. Mean values, summed for all populations, are $F_{IS}=0.192$, $F_{IT}=0.299$ and $F_{ST}=0.133$, which indicate that there is a limited degree of interpopulational differentiation with most variance residing within populations.

Polymorphism over 17 loci and six populations is 56.8%. Heterozygosity averaged over the subset of the 12 polymorphic loci is 26.9% (Table 3). Heterozygotes range from 11.9% for

TABLE 1 Seed sets in *Dedeckera eurekaensis*

Population*	Collection year	% Developed (filled)	%Developed (aborted)	% Partially developed	% Non-developed	n (ovaries)	n (plants)
Bishop†	1988	1.6	6.5	39.9	52.0	1,624	10
	1987	0.6	20.2	72.5	5.8	1,046	10
	1984	4.6	50.1	24.5	20.8	351	1
Coldwater Cyn‡	1988	4.7	25.8	67.4	2.0	2,875	10
	1987	4.2	48.7	44.3	2.8	1,112	10
	1985	1.9	33.3	44.8	20.0	105	1
Gunter Cyn§	1988	0.8	41.3	32.9	25.0	1,482	10
	1987	1.9	54.1	36.2	7.8	1,238	10
Dedeckera Cyn	1988 (upper)	4.0	29.0	65.1	1.7	1,682	10
	1987	1.7	67.3	28.5	2.4	1,230	10
	1988 (lower)*	4.3	44.3	49.6	1.8	1,464	10
	1987	2.9	61.3	33.2	2.6	1,265	11
	1978	1.7	46.0	42.5	9.7	113	1
	1975	2.9	29.7	50.5	16.8	101	1
Last Chance Mts#	1987	1.9	70.4	26.7	1.0	1,184	10
Panamint Mts**	1980	0.9	49.5	26.6	22.9	109	1
Means or totals (overall)		2.5	42.3	42.8	12.2	16,981	116
Means or totals	1988	3.1	29.4	51.0	16.5	9,127	50
Means or totals	1987	2.2	54.6	39.5	3.7	7,075	61
Means or totals	1975-85	2.4	44.5	33.5	18.9	779	5

The structures scored are anatomically indehiscent, single-seeded, post-anthesis ovaries or maturing fruits (achenes), although functionally they behave essentially as seeds. Developed (aborted) fruits reach approximately full length, but the ovary wall is partially collapsed between the ribs and the ovule has atrophied. Partially developed fruits did not reach full size, but all initiated ovary expansion before ovule abortion, which is taken as *prima facie* evidence for fertilization. Non-developed ovaries showed no evidence of expansion. They were probably not pollinated, or the ovules were unfertilized, or abortion occurred at the earliest stages of embryogenesis or endosperm development.

* All locations on Universal Transverse Mercator grid ticks, zone 11; Cyn, Canyon; Mts, mountains. † 385500E-4136800N; ‡ 384000E-4149000N; § 388100E-4146200N; || 443700E-4101900N; ¶ 443600E-4101300N; # 442000E-4094800N; ** 434400E-4083100N.

TABLE 2 Controlled pollinations in *Dedeckera eurekaensis*

Pollination treatment	Per cent seed types recovered				N 1-ovuled (ovaries)	N (plants)
	Developed (filled)	Developed (aborted)	Partially developed	Non-developed		
Intrapopulation crosses	11.4	21.1	66.7	0.88	228	10
Self-pollinations (geitonogamous)	6.3	7.9	85.7	0	63	4
Interpopulation crosses ($\times$ Gunter Cyn.)	12.9	37.1	48.4	1.6	62	2
Spontaneous seed set in parental plants	4.7	25.8	67.4	2.0	2,875	10
Flowers isolated in nylon enclosures	3.9	26.0	65.5	4.6	281	3

Coldwater Canyon population, 2-29 July 1988 (see Table 1 legend for locality and explanation of fruit types). Control flowers were taken from unmanipulated portions of experimental plants and left open so that pollination was spontaneous. Isolated flowers were those simply left under nylon enclosures (after removing all open or fruiting flowers), thus excluding all insect pollen vectors. All open or fruiting flowers were removed from the experimental inflorescences after which they were enclosed with removable fine-mesh nylon covers mounted over a wire frame to prevent contact of the nylon with the flowers. Flowers were emasculated daily before anthesis. All three stigmas were pollinated by contacting their sticky surfaces with a dehiscing anther after the styles had diverged and the stigmas became receptive. The small flowers (2-3 mm) required that all floral manipulations be effected under a dissecting microscope mounted on a camera tripod. The nylon covers were removed after the completion of all pollinations.

TABLE 3 Genetic variability at 12 loci

	Mean sample size per locus	Mean no. of alleles per locus	% Loci poly- morphic	Mean hetero- zygosity
White Mountains populations				
Bishop	15.8	1.9	66.7	0.211
Coldwater Cyn	18.8	2.6	75.0	0.218
Gunter Cyn	16.3	2.3	75.0	0.194
Means	17.0	2.3	72.2	0.207
Last Chance Mountains populations				
Dedeckera Cyn (upper)	14.8	2.7	91.7	0.384
Dedeckera Cyn (lower)	16.7	3.3	91.7	0.293
Last Chance Mts	13.3	2.7	83.3	0.318
Means	14.9	2.9	88.9	0.332

Values in this table are derived from analysis of 12 loci—EST-2 (esterase), GSR-2 (glutathione reductase), IDH-2 (isocitrate dehydrogenase), MDH-1, MDH-2, MDH-4 (malate dehydrogenase), MDR-1 (menadiol reductase), GPI-2 (glucose phosphate isomerase), PGM-1, PGM-2, PGM-3 (phosphoglucose mutase) and SKDH (shikimate dehydrogenase)—that were polymorphic in at least one of the six populations. When five additional loci ADK-1 (adenylate kinase), GPI-1, GSR-1, LAP (leucine amino peptidase), and MDH-3, monomorphic for all populations, are included, the mean polymorphism is 56.8%. A locus was considered polymorphic if any allelic variant was detected. Mean proportion of heterozygotes (direct count) is less than the expected Hardy-Weinberg (unbiased estimate)²⁴. The fixation index F_{25} showed that only two loci (PGM-1 in the Gunter Cyn population and SKDH in the population of the Last Chance Mountains) had significant ($P < 0.001$) deviations in genotype proportions.

MDR-1 to 58.3% for GPI-2. These heterozygosity averages are underestimated because they include the zero values for those markers that are not polymorphic in the particular population sampled. This amount of variation is much higher than that reported for other species of desert shrubs¹⁶, but is similar to other woody, outcrossing perennials¹⁷⁻¹⁸.

Dedeckera is vegetatively vigorous and flowers profusely. Its longevity (at least 140 years and probably much more) may be correlated with heterozygosity as reported in *Liatris*¹⁹ (Asteraceae). In *Dedeckera*, vegetative fitness may be largely dependent upon rare, highly heterotic (or possibly epistatic) genotypes. The uniquely heterozygous genotypes that do survive may, however, have low reproductive potential because of the excessively high segregational genetic load. Reproduction could be further compromised by the accumulation of recessive lethals in long-lived meristems²⁰ and chromosomal mutations²¹.

The persistence of palaeoendemics may depend on rare multiple-locus heterotic genotypes that could arise as a result of tracking some protracted secular environmental change, for example, increasing aridity. The pace and extent of such change could easily exhaust the species' additive genetic variance relevant to adapting to that change. Heterosis may be the only survival strategy available to such species, in spite of the negative reproductive consequences. Thus the 13% increase in mean heterozygosity for the populations of *Dedeckera* in the Last Chance Mountains, as opposed to those in the White Mountains (Table 3), may be important because the former localities are both hotter and drier. In fact, many palaeoendemics may be ecologically 'out of place', in that they may not possess many of the adaptations typical of plants occupying that habitat—thus *Dedeckera* flowers in mid-summer when desert perennials are typically dormant.

Rabbits²² and humans²³ have genetically mediated spontaneous embryo abortion rates of about 50% and 70%, respectively. The loss of reproductive capacity stemming from early embryonic genetic load should not be overlooked as a possible element in the decline of higher animal palaeoendemics. Such cases may be observed rarely because higher animals generally have much shorter life-spans, relatively limited reproductive potentials, and because early abortion is more difficult to detect than in plants.

Other palaeoendemic plants in North America, Africa and Australia also seem to have exceedingly low seed sets. Reproductive capacity should be given careful consideration in management decisions regarding rare and/or endangered species. □

Received 8 November 1988; accepted 16 January 1989.

- Wiens, D. *Oecologia* **64**, 47–53 (1984).
- Harper, J. L. & Wallace, H. L. *Oecologia* **74**, 31–38 (1987).
- Wiens, D., Calvin, C., Wilson, C. A., Frank, D. & Seavey, S. *Oecologia* **71**, 501–509 (1987).
- Bonnell, M. L. & Selander, R. K. *Science* **184**, 908–909 (1974).
- O'Brien, S. J. *et al. Science* **227**, 1428–1434 (1985).
- Wildt, D. E. *et al. Nature* **329**, 328–331 (1987).
- Ledig, F. T. & Conkle, M. T. *Evolution* **37**, 79–85 (1983).
- Waller, D., O'Malley, D. & Gawler, S. *Cons. Biol.* **1**, 335–340 (1987).
- Meinke, D. W. & Sussex, I. M. *Dev. Biol.* **72**, 50–61 (1979).
- Marsden, M. P. F. & Meinke, D. W. *Am. J. Bot.* **72**, 1801–1812 (1985).
- Raghaven, V. *Embryogenesis in Angiosperms* (Cambridge University Press, New York, 1986).
- Steenbergh, W. F. & Lowe, C. H. *Ecology* **50**, 825–834 (1969).
- Turner, R. M., Alcorn, S. M. & Olin, G. *Ecology* **50**, 835–844 (1969).
- Wiens, D., DeDecker, M. & Wiens, C. D. *Madroño* **33**, 302–305 (1986).
- Nei, M. *Ann. hum. Genet.* **41**, 225–233 (1977).
- Hamrick, J. L., Mitton, J. B. & Linhart, Y. B. *Proc. Symp. Isozymes of North American Forest Trees and Forest Insects* (ed. Conkle, M. T.) 35–41 (USDA, Forest Service Gen. Tech. Rep. PSW-48, 1981).
- Ledig, F. T. *Cons. Biol.* (ed. Soule, M. E.) 77–104 (Sinauer, Sunderland, Massachusetts, 1986).
- Gottlieb, L. D. *Progress in Phytochemistry* (eds Rheinhold, L., Harborne, J. B. & Swain, T.) 1–40 (Pergamon, New York, 1981).
- Schaal, B. & Levin, D. A. *Am. Nat.* **110**, 191–206 (1977).
- Klekowski, E. Jr, Mohr, H. & Kazarinova-Fukshansky, N. *Genetics, Development, and Evolution* (eds Gustafson, J. P., Stebbins, G. L. & Ayala, F. J.) 79–113 (Plenum, New York, 1986).
- Sankaranarayanan, K. *Mutat. Res.* **61**, 1–26 (1979).
- Brambell, F. W. R. *Biol. Rev.* **23**, 370–407 (1948).
- Biggers, J. D. *New Engl. J. Med.* **304**, 336–342 (1981).
- Nei, M. *Genetics* **89**, 583–590 (1978).
- Wright, S. *Evolution and Genetics of Populations* (University of Chicago Press, 1978).

ACKNOWLEDGEMENTS. We thank J. Chesnut, J. Neuhauser, M. Kobler, S. Wroe, W. Owen, O. Pollak, C. Scheidlinger, S. Hodges and M. Weiss for field and laboratory assistance, and the staff of the White Mountain Research Station, Bishop, California, for laboratory facilities. J. Ehleringer, T. Hazelrigg, D. Mansfield, K. Paige, P. Raven, R. K. Vickery Jr, E. O. Wilson, and especially J. Endler, have read and commented on the manuscript. M. DeDecker, J. Dickinson and W. Wiens offered various suggestions. J. Rourke first mentioned the existence of plants with exceptionally low sexual reproductive capacity in Africa. The study was supported by grants from the NSF, the USDA (Forest Service, Rocky Mountain Forest and Range Experiment Station) and the University of Utah, College of Science.

The *ninaA* gene required for visual transduction in *Drosophila* encodes a homologue of cyclosporin A-binding protein

Bih-Hwa Shieh, Mark A. Starnes, Suzanne Seavello, Greg L. Harris & Charles S. Zuker

Department of Biology and Center for Molecular Genetics, University of California, San Diego, La Jolla, California 92093, USA

MUTATIONS of the *Drosophila melanogaster ninaA* gene affect phototransduction: *ninaA* mutant flies have a 10-fold reduction in the levels of rhodopsin in the R1–R6 photoreceptor cells^{1,2}. The *ninaA* gene was isolated and found to encode a 237-amino-acid protein that has over 40% amino-acid sequence identity with the vertebrate cyclosporin A-binding protein, cyclophilin, a protein that seems to be involved in T-lymphocyte activation. The remarkable evolutionary conservation of cyclophilin in two phylogenetically distant organisms and its involvement in diverse transduction processes suggests that this protein plays an important role in cellular metabolism. Indeed, cyclophilin has recently been shown to be a prolyl *cis-trans* isomerase that catalyses, *in vitro*, rate-limiting steps in the folding of a number of proteins^{3,23}. Here, we present evidence for the involvement of cyclophilin-like molecules in a defined cellular process. The availability of mutations in a cyclophilin gene provides a new model system for the study of cyclophilin and cyclosporin action.

Numerous mutations that affect visual transduction in *Drosophila* have been isolated (reviewed in refs 4, 5). To isolate genes encoding phototransduction-specific proteins, we used a subtractive hybridization protocol to identify genomic clones

containing sequences preferentially expressed in the adult visual system. Clones that tested positive in our screen (see Fig. 2 legend) were isolated and used in *in situ* hybridizations to polytene chromosomes to determine their cytogenetic map positions. The *ninaA* locus has been mapped to the second chromosome, position 21D3-E2 (ref. 1). Two of the 120 genomic clones we isolated, λ 322 and λ N4, were mapped to position 21D4-E2, within the cytogenetic map location of *ninaA*. The *ninaA* locus constitutes one of eight complementation groups of *Drosophila* mutants with drastically reduced rhodopsin levels^{1,2}. *ninaA* flies have a severe reduction of rhodopsin (Rh1) levels in the R1-R6 photoreceptor cells (Fig. 1). The reduction of rhodopsin levels in the R1-R6 photoreceptors of *ninaA* mutant flies has recently been shown⁶ not to be due to reduced expression of the R1-R6 opsin gene (*ninaE*) but may be the result of a defect in some aspect of post-translational processing of the *ninaE* opsin, or a secondary effect on the stability of rhodopsin due to a defect in phototransduction.

Characterization of clones λ 322 and λ N4 showed that they contain overlapping sequences and encode a 0.9 kilobase (kb) RNA (Fig. 2) that is present in preparations of poly (A)⁺ RNA from wild-type heads, but not from wild-type bodies or heads from mutant flies lacking the compound eyes (*eya*; eyes absent⁷) (Fig. 2b). We used the large *Eco*RI restriction fragment of λ 322 to screen a *Drosophila* head cDNA library and isolated several cDNA clones. Using M13 dideoxynucleoside triphosphate sequencing we have determined the DNA sequence of one of those cDNA clones and of the 1.2 kb *Hind*III-*Pst*I genomic fragment of λ 322 (Fig. 2a). Figure 3a shows the nucleotide sequence and the deduced amino-acid sequence of the *ninaA* gene product. The structure of the RNA (Fig. 2a) was determined by analysing genomic and cDNA sequences. Confirmation that λ 322 includes the *ninaA* gene was obtained by determining the nucleotide sequence of the *ninaA*^{P228} mutant allele. Transcript size and levels are normal in these mutant flies (Fig. 2b). The mutant allele was isolated and cloned by a polymerase chain reaction. The *ninaA*^{P228} allele has all the nucleotide sequence polymorphisms of the Oregon-R strain. Interestingly, however, this allele has a single coding nucleotide change at position 895, from TGG to TGA, causing a change from the encoded tryptophan to an opal translation termination codon (Fig. 3a arrow). The nature of this change is consistent with the chemical mutagenic origin of *ninaA*^{P228} (ref. 1). Thus, the cytogenetic location, expression profile, and altered nucleotide sequence in mutant flies are all consistent with this being the *ninaA* gene.

Comparison of the deduced amino-acid sequence of *ninaA* with previously sequenced genes and proteins revealed very high homology to the bovine cyclosporin A-binding protein cyclophilin (Fig. 3b). These two proteins have more than 40% amino-acid identity throughout their length (shaded boxes); conserved

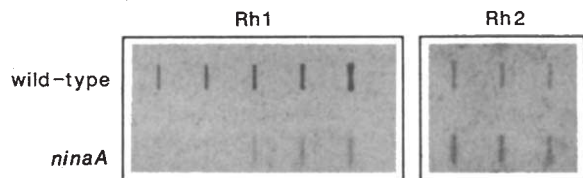


FIG. 1 *ninaA* mutants have a severe reduction of Rh1 opsin levels. Protein extracts were prepared from the heads of control and *ninaA* flies by digitonin extraction²³. Serial dilutions (0.25, 0.5, 1, 2 and 4 μ g) were slot-blotted onto nitrocellulose paper and processed for western analysis using a monoclonal antibody against the Rh1 opsin (kindly provided by Dr H. Gert de Couet). Mutant flies have a 8–10-fold reduction of the opsin present in the R1–R6 photoreceptors (Rh1). The right panel shows similar extracts (4, 2 and 1 μ g) treated with a monoclonal antibody against the ocellar (Rh2) opsin. This monoclonal antibody was generated against a Rh2-specific synthetic peptide. Note the specificity of the mutant phenotype for the Rh1 opsin (see also refs 1, 2).

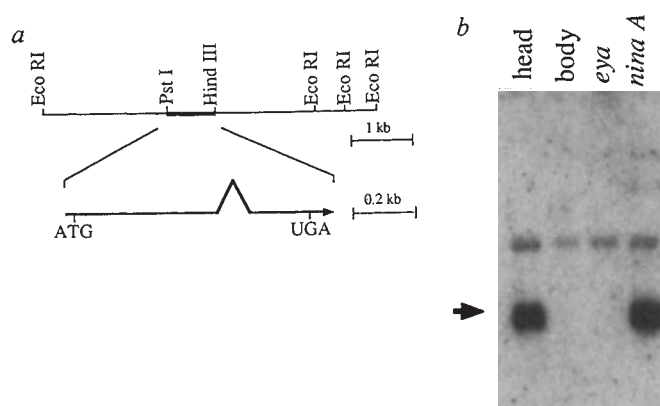


FIG. 2 λ 322 encodes an eye-specific transcript. a, Restriction map of λ 322 and structure of the RNA it encodes. A map of λ 322 indicating the restriction sites for *Hind*III, *Pst*I and *Eco*RI (top map). The diagram below the map shows the structure of the RNA as deduced by comparison of the nucleotide sequence of a cDNA clone and the genomic sequence. b, Poly (A)⁺ RNAs were extracted from adult heads of wild-type, *eya* and *ninaA* flies and from adult bodies. The RNAs (3 μ g per lane) were gel-fractionated, blotted, and hybridized to a 1.2-kb λ 322 radiolabelled *Pst*I-*Hind*III fragment. λ 322 hybridizes to a 0.9-kb eye-specific transcript (arrow). The blot was also hybridized to a probe encoding 'common' RNAs to control for the amount of RNA loaded (λ 9, unpublished observations). Hybridization to λ 9 mRNA is seen as the band present in all lanes with a mobility significantly slower than λ 322 mRNA. An RNA ladder (BRL) was used to provide size markers.

METHODS. λ 322 was isolated by hybridization with eye-specific probes. Poly (A)⁺ RNA, isolated from the heads of wild-type flies, was used to template the synthesis of cDNA. This cDNA was hybridized in solution to a 20-fold excess of poly(A)⁺ RNA isolated from the bodies of adult flies. Single-stranded molecules were then separated from the double-stranded DNA-RNA heteroduplexes by hydroxylapatite chromatography²⁴. Hybridizations were carried out in 30 μ l of 0.5 M phosphate buffer (pH 6.9) containing 2 μ g of head cDNA and 40 μ g of body RNA. The reaction mixtures were incubated for 24 h at 65 °C. The resulting single-stranded cDNA, representing mostly 'head-specific' sequences, were used in a second round of subtraction with RNA isolated from heads of flies carrying the *eya* mutation⁷. Single-stranded molecules were fractionated by hydroxylapatite chromatography and used as templates for the synthesis of very high specific activity second-strand cDNA by two rounds of amplification with random primers (pN₆, Amersham). This radiolabelled cDNA, representing a probe highly enriched in eye-specific sequences, was then used to screen a genomic library. Subtractions with *eya* head poly(A)⁺ RNA were carried out with 0.2 μ g of head-enriched cDNA and 4 μ g of *eya* RNA. About 10% of input cDNA was recovered after the first subtraction, and about 20% after the second. First and second-strand cDNA synthesis were carried out as described²⁵.

substitutions account for over 30% of the remaining amino-acid residues. Cyclophilin is a soluble protein with a relative molecular mass (M_r) of 17,000 with high binding affinity and stereospecificity for cyclosporin A (CsA)^{8–11}. CsA is an 11-amino-acid cyclic peptide of fungal origin with potent immunosuppressive properties widely used to prevent graft rejection and for the treatment of autoimmune disorders (reviewed in refs 12–14). Recent data support the hypothesis that the action of CsA is mediated through cyclophilin^{15–17}. Although the exact mechanism of CsA action is not known, there is good evidence from whole cell systems that CsA interrupts an early event in the antigenic activation of T helper cells that results in lymphokine production^{18–20}.

To determine whether *Drosophila* does indeed contain proteins with CsA-binding activity, we carried out [³H]CsA binding assays on head extracts prepared from wild-type controls, *eya*, and *ninaA*^{P228} flies. As shown in Fig. 4, head extracts have significant levels of cyclophilin-like activity (specific partition on sepharose LH-20 chromatography^{8–11}); about 30% of the head binding is associated with the visual system in that it is removed by the *eya* mutation. The *ninaA*^{P228} mutation leads to a similar decrease in CsA binding indicating that the *ninaA*

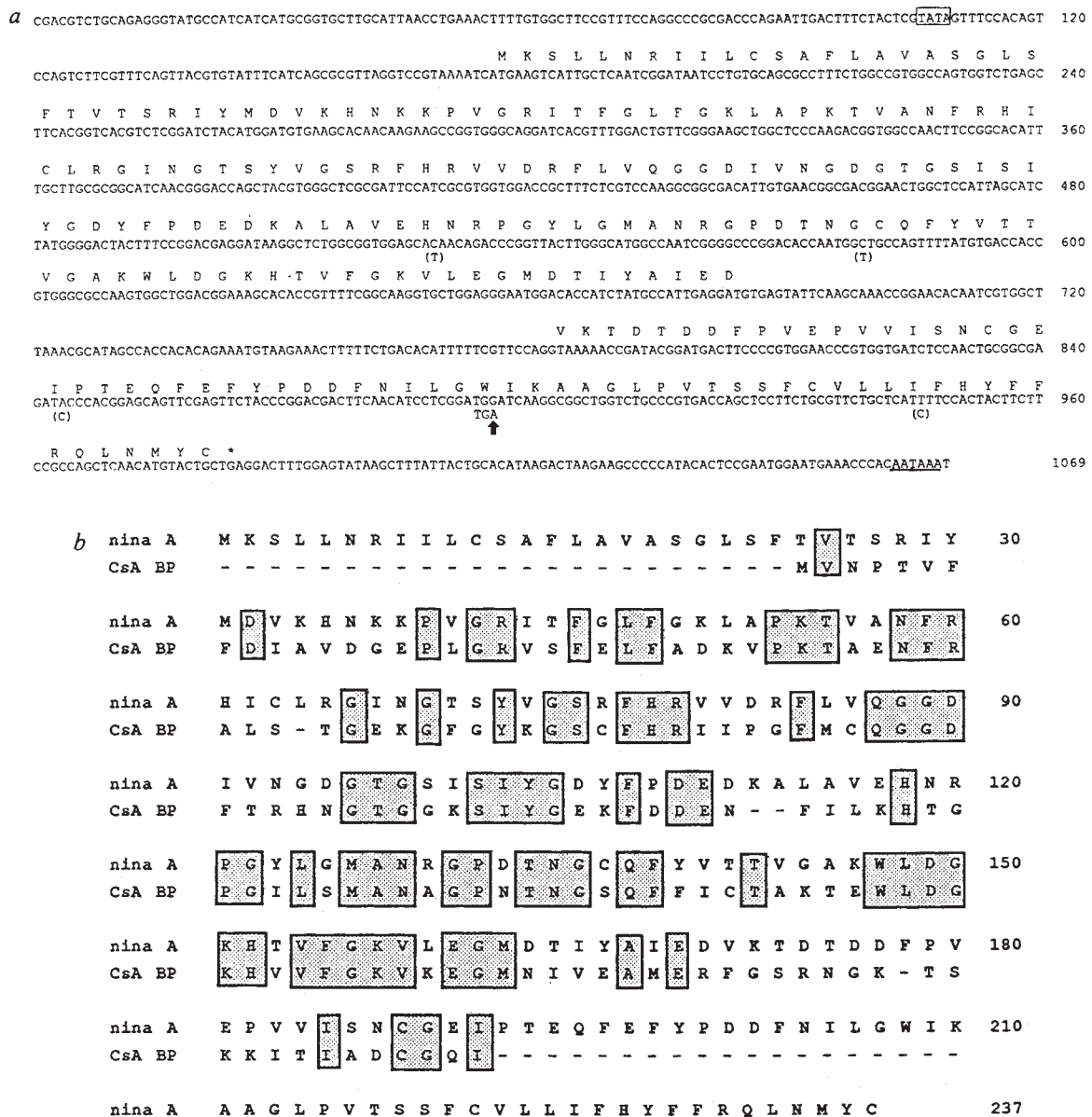


FIG. 3 Nucleotide sequence and deduced amino-acid sequence of the *ninaA* gene. **a**, The sequence shown was determined on both strands of the genomic and cDNA clones²⁶. The boxed region at positions 106–110 shows the putative TATA box. The putative polyadenylation signal is underlined at positions 1,063–1,068. The deduced protein sequence is shown aligned under the nucleotide sequence. The initiator methionine was assigned as the first in-frame methionine in the sequence. The gap at nucleotides 686–782 indicate the presence of a 96-base pair (bp) intron. The arrow at position 895 highlights the single nucleotide change in the *ninaA*^{F228} mutant gene (G→A). This change creates a translation termination codon. The single nucleotides below positions 528, 579, 844 and 946 show the differences in the coding regions between the cDNA (Oregon R, P2 strain) and the genomic (Canton S) sequence; none of these polymorphisms would result in a change in the encoded amino

acid. **b**, A co-linear alignment of the deduced amino-acid sequence of the *ninaA* gene and the bovine cyclophilin protein¹⁰. Cyclophilin from mammalian species analysed so far have over 95% sequence identity, and the physical properties of the CsA binding proteins from other vertebrate species are remarkably similar^{10,27,28}. Amino acids are designated by the single-letter code. The alignment has been optimized for the largest number of identities with the least number of gaps. Boxed areas indicate the amino-acid identities between the two proteins.

METHODS. The mutant allele was cloned by selective amplification of the *ninaA* gene by polymerase chain reaction using protocol and reagents supplied by Perkin-Elmer Inc. (PCR kit N801-0043). Oligonucleotides complementary to positions –8 to +11 and +986 to +1,005 were used as primers.

locus encodes a cyclophilin-like protein that accounts for most of the visual system-specific CsA-binding activity (Fig. 4). The binding activity remaining in the *eya* and *ninaA* extracts probably represents the presence of non-eye-specific cyclophilin-like proteins.

The *ninaA* R1–R6 photoreceptors have about 10% of wild-type levels of rhodopsin^{1,2}. Zuker and co-workers⁶ have previously shown that the Rh1 opsin gene is expressed at normal levels in *ninaA* mutants. The reduction of rhodopsin must therefore be a post-transcriptional or a post-translational event. The discovery that cyclophilin encodes a prolyl *cis-trans* isomerase³

suggests that it may be required for the correct folding of proline-containing polypeptides. We propose that this isomerase activity is necessary for the correct folding and the stability of rhodopsin in *Drosophila* R1–R6 photoreceptors. The requirement for large amounts of rhodopsin in these cells may explain the existence of a photoreceptor cell-specific form of this enzyme. Moreover, the existence of a cell-specific isoform suggests that there are specific substrates for these proteins and explains the presumed diversity of cyclophilin-like molecules (M.A.S. and C.S.Z., unpublished observations and refs 3, 10). It is therefore possible that, in the immune system, CsA blocks

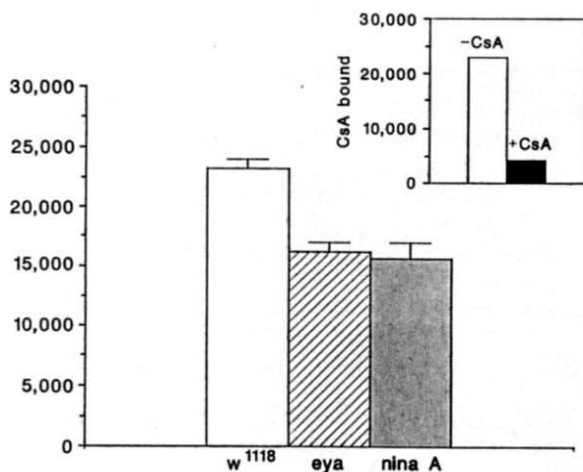


FIG. 4 *ninaA* flies have reduced levels of cyclosporin A-binding activity. Extracts prepared from heads of control *w¹¹¹⁸* flies, *eya* or *ninaA* flies were tested for CsA-binding activity⁸. The results indicate that *Drosophila* head extracts contain significant amounts of cyclophilin-like activity (approximately 30 ng CsA binding mg^{-1} extract) and that all eye-associated binding is reduced in *ninaA* flies (compare *eya* and *ninaA* extracts). Bars above the graph indicate standard errors (*w¹¹¹⁸*, *n*=10; *eya*, *n*=8; *ninaA^{P228}*, *n*=7). *Drosophila* extracts were prepared from heads of wild-type or mutant individuals exactly as described⁸. Binding assays were carried out with ³H-cyclosporin A (17 Ci mmol^{-1} Amersham) at either 30 °C or 37 °C and the products separated by partition chromatography on a Sephadex LH-20 column⁸. The inset shows the specificity of the binding assay as determined by competition with unlabelled CsA.

the activity of a cyclophilin required for the proper functioning of the antigen-mediated transduction pathway.

In the visual system, the interconversion between the active (metarhodopsin) and inactive (rhodopsin) states of the visual pigment molecule involves significant conformational changes. In the invertebrate visual cascade, these two forms are thermally stable and photoconvertible. It would therefore be very interesting to determine whether the isomerase encoded by *ninaA* is important in this event during the transduction cycle. The availability of *Drosophila* lines carrying mutations in the endogenous *ninaA* gene, and the use of P-element-mediated germline transformations^{21,22}, may allow for the functional expression of wild-type and modified alleles in their normal cellular and organismal environment. A combined biochemical, physiological and molecular genetic dissection will help assign specific roles to the *ninaA* gene product in the phototransduction process. □

Received 22 November 1988; accepted 6 January 1989.

- Larrievée, D. C., Conrad, S., Stephenson, R. S. & Pak, W. L. *J. gen. Physiol.* **78**, 521–545 (1981).
- Stephenson, R. S., O'Tousa, J., Scavarda, N. J., Randall, L. L. & Pak, W. L. in *The Biology of Photoreception* (eds Cosens, D. J. & Vince-Price, D.), 477–501. (Cambridge University Press, 1983).
- Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T. & Schmid, F. X. *Nature* **337**, 476–481 (1989).
- Hall, J. C. Q. *Rev. Biophys.* **15**, 223–479 (1982).
- Heisenberg, M. & Wolf, R. in *Vision in Drosophila: Genetics of Microbehavior, Studies of Brain Function Vol. 12* (ed. Braisted, V.). (Springer, Berlin, 1984).
- Zuker, C. S., Mismar, D., Hardy, R. & Rubin, G. M. *Cell* **53**, 475–485 (1988).
- Sved, J. *Drosoph. Info. Serv.* **73**, 169 (1986).
- Handschumacher, R. E., Harding, M. W., Rice, J., Druggie, R. J. & Speicher, D. W. *Science* **226**, 544–547 (1984).
- Merker, M. M. & Handschumacher, R. E. *J. Immun.* **132**, 3064–3070 (1984).
- Harding, M. H., Handschumacher, R. E. & Speicher, D. W. *J. biol. Chem.* **261**, 8547–8555 (1986).
- Koletsky, A. J., Harding, M. W. & Handschumacher, R. E. *J. Immun.* **137**, 1054–1059 (1986).
- Borel, J. F., Feurer, C., Gubler, H. & Stahelin, H. *Agents and Actions* **6**, 468–475 (1976).
- Cohen, D. J., et al. *Ann. intern. Med.* **101**, 667–682 (1984).
- Shevach, E. M. A. *Rev. Immun.* **3**, 397–423 (1985).
- Druggie, R. J. & Handschumacher, R. E. *Transplant Proc.* **20**, Suppl. 2, 301–309 (1988).
- Second International Congress on Cyclosporine, *Transplant Proc.* **20**, Suppl. 2 (1988).
- International Symposium on the Mechanism of Action of Cyclosporine Transplantation **46**, 15 (1988).
- Manger, B., Hardy, K. J., Weiss, A. & Stobo, J. D. *J. clin. Invest.* **77**, 1501–1506 (1986).
- Szamel, M., Berger, P. & Resch, K. *J. Immun.* **136**, 264–269 (1986).
- Rosoff, P. M. & Terres, G. *J. Cell Biol.* **103**, 457–463 (1986).
- Rubin, G. M. *Trends Neurosci.* **8**, 231–233 (1985).

- Rubin, G. M. *Science* **240**, 1452–1459 (1988).
- Ostrov, S. E., Wilson, M. & Pak, W. L. *Biochem. biophys. Res. Commun.* **59**, 960–966 (1974).
- Davis, M. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2194–2198 (1984).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. in *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Haendler, B., Hofer-Warbinek, R. & Hofer, E. *EMBO J.* **6**, 947–950 (1987).
- Danielson, P. E. et al. *DNA T.* **261**–267 (1988).
- Takahashi, N., Hayano, T. & Suzuki, M. *Nature* **337**, 473–475 (1989).

ACKNOWLEDGEMENTS. We thank G. Fischer and co-workers for communicating their results before publication. We thank W. A. Harris for comments, suggestions and helpful discussions. We also thank G. M. Rubin, R. E. Handschumacher, J. Hall, J. Posakony, M. Montal and members of this laboratory for critical reading of the manuscript. We particularly want to thank G. Kline for technical assistance and K. Blumeyer for her help on polymerase chain reactions. Dr W. Pak kindly provided the *ninaA^{P228}* stock. This work was supported by grants from the NIH to C.S.Z. and a grant-in-aid from the Fight for Sight foundation to B.-H. S. C.S.Z. acknowledges support from the McKnight endowment Fund for Neuroscience and the Pew Scholars Program in the Biomedical Sciences.

Identification of a photoreceptor-specific mRNA encoded by the gene responsible for retinal degeneration slow (*rds*)

Gabriel H. Travis, Miles B. Brennan,
Patria E. Danielson, Christine A. Kozak*
& J. Gregor Sutcliffe

Research Institute of Scripps Clinic, 10666 N. Torrey Pines Road,
Department of Molecular Biology, La Jolla, California 92037, USA

* Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, NIH, Bethesda, Maryland 20892, USA

MUTANT mice homozygous for 'retinal degeneration slow' (*rds/rds*) are characterized phenotypically by abnormal development of photoreceptor outer segments in the retina, followed by gradual degeneration of photoreceptors^{1–3}. This process of degeneration is complete by one year, with preservation of all other retinal cells⁴. The biochemical defect that leads to the mutant phenotype is not known. Our strategy for cloning the *rds* gene was based upon three previously reported observations. First, the *rds* locus maps to chromosome 17^{5,6}. Second, experimental *rds/rds* ↔ +/+ and *rds/+* ↔ +/+ tetra-parental mice manifest patchy photoreceptor changes in the retina^{7,8}, suggesting that the wild-type *rds* locus is expressed within cells of the photoreceptor lineage. Finally, the process of degeneration is specific to photoreceptors. On the basis of these observations, we predicted that the *rds* mRNA is encoded by a gene on chromosome 17 and is normally expressed exclusively within photoreceptors in the retina. We here present evidence that this is the case.

Given our predictions, a cDNA representing a photoreceptor-specific mRNA encoded by a gene on chromosome 17 would be a candidate clone of the *rds* mRNA. To isolate cDNA clones of photoreceptor-specific mRNAs, we took advantage of the unrelated mouse mutant, retinal degeneration (*rd/rd*)^{9,10}. Mice homozygous for this mutation manifest rapid degeneration of photoreceptors, a process that is virtually complete by four weeks, with the preservation of all other retinal cell types^{11,12}. Therefore, an mRNA present in wild-type (C57BL/6) but absent from fully degenerate *rd/rd* (C3H/HeJ) retina is photoreceptor-specific. cDNA clones of twelve different photoreceptor-specific mRNAs were isolated from an adult C57BL/6 mouse retina library by subtractive and differential colony screening¹³ of 6–7-week-old C57BL/6 minus 6–7-week-old C3H/HeJ retina. Northern blot hybridization patterns for six of these clones are shown in Fig. 1a.

The chromosome assignments for the genes encoding each of the 12 photoreceptor-specific mRNAs were made by probing a panel of mouse × hamster hybrid cell-line DNAs¹⁴ with a representative clone of each of the photoreceptor-specific mRNAs. Clone IG3 mapped to chromosome 17 with 100% concordance

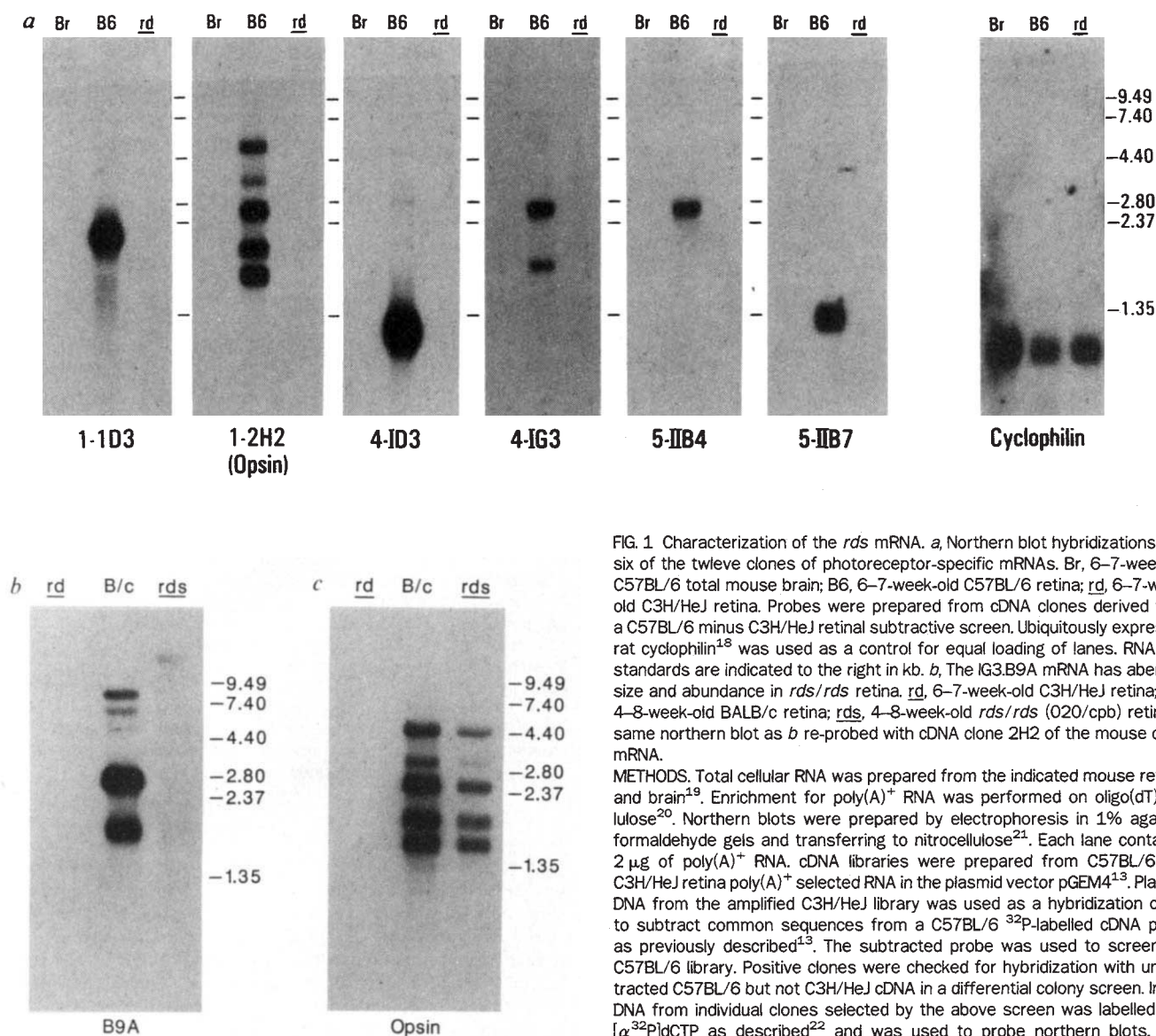


FIG. 1 Characterization of the *rds* mRNA. *a*, Northern blot hybridizations with six of the twelve clones of photoreceptor-specific mRNAs. Br, 6-7-week-old C57BL/6 total mouse brain; B6, 6-7-week-old C57BL/6 retina; rd, 6-7-week-old C3H/HeJ retina. Probes were prepared from cDNA clones derived from a C57BL/6 minus C3H/HeJ retinal subtractive screen. Ubiquitously expressed rat cyclophilin¹⁸ was used as a control for equal loading of lanes. RNA size standards are indicated to the right in kb. *b*, The IG3.B9A mRNA has aberrant size and abundance in *rds/rds* retina. rd, 6-7-week-old C3H/HeJ retina; B/c, 4-8-week-old BALB/c retina; rds, 4-8-week-old *rds/rds* (O20/cpb) retina. *c*, same northern blot as *b* re-probed with cDNA clone 2H2 of the mouse opsin mRNA.

METHODS. Total cellular RNA was prepared from the indicated mouse retinas and brain¹⁹. Enrichment for poly(A)⁺ RNA was performed on oligo(dT) cellulose²⁰. Northern blots were prepared by electrophoresis in 1% agarose formaldehyde gels and transferring to nitrocellulose²¹. Each lane contained 2 µg of poly(A)⁺ RNA. cDNA libraries were prepared from C57BL/6 and C3H/HeJ retina poly(A)⁺ selected RNA in the plasmid vector pGEM4¹³. Plasmid DNA from the amplified C3H/HeJ library was used as a hybridization driver to subtract common sequences from a C57BL/6 ³²P-labelled cDNA probe as previously described¹³. The subtracted probe was used to screen the C57BL/6 library. Positive clones were checked for hybridization with unsubtracted C57BL/6 but not C3H/HeJ cDNA in a differential colony screen. Insert DNA from individual clones selected by the above screen was labelled with [³²P]dCTP as described²² and was used to probe northern blots. Final stringency wash conditions were 0.2 × SSC/0.2% SDS at 68 °C.

(Table 1). A full-length cDNA clone (IG3.B9A), containing an insert of approximately 2.7 kilobase (kb), was isolated by re-probing the C57BL/6 retinal cDNA library with clone IG3.

Clone IG3.B9A was used to probe northern blots containing RNA from the retinas of 1-2-month-old (pre-degenerate) *rds/rds* mice (BALB/c background) and age-matched wild-type (BALB/c) mice (Fig. 1*b*). Two major mRNA species of approximately 1.6 kb and 2.7 kb (estimated abundance, 0.05%) were observed in the BALB/c retina lane. These bands were not present in the *rds/rds* retina lane. A faint doublet band of approximately 12 kb was detected in the *rds/rds* lane however, which was not present in the BALB/c lane. Thus, the mRNAs detected with clone IG3.B9A in the retinas of *rds* mice are aberrant in size and abundance, suggesting that IG3.B9A is a clone of the wild-type *rds* mRNA.

When the same northern blot was re-probed with a cDNA clone of the opsin mRNA (clone 1-2H2 in Fig. 1*a*), the BALB/c and *rds/rds* retina lanes exhibited identical hybridization patterns, although the signal intensity in the mutant lane was slightly reduced (Fig. 1*c*), consistent with partial degeneration of photoreceptors. This experiment excluded the possibility that the observed difference in RNA hybridization patterns with the

IG3.B9A probe in wild-type and *rds/rds* retina was due to premature degeneration of photoreceptors in the mutant, or to a generalized abnormality in RNA processing within photoreceptors in *rds/rds*.

To define the site of the mutation within the putative *rds* locus, five adjacent non-overlapping restriction endonuclease fragments spanning the full-length cDNA clone IG3.B9A were prepared (Fig. 2*a*). These fragments were used individually as probes to examine restriction digests of genomic DNA prepared from BALB/c and *rds/rds* mice on Southern blots. Restriction fragment length polymorphisms (RFLPs) were detected in the DNA of BALB/c relative to *rds/rds* mice with probe B9A-II (Fig. 2*b*), but not with probes B9A-I, -III, -IV or -V (data not shown).

Genomic libraries prepared with the DNA from BALB/c and *rds/rds* mice were probed with cDNA clone IG3.B9A. A clone B7 was isolated from the BALB/c library and a clone R24 was isolated from the *rds/rds* library. The λ-cloned DNAs gave similar hybridization patterns to their corresponding wild-type or mutant genomic DNAs on Southern blots probed with cDNA clone fragment B9A-II (Fig. 2*b*).

To define the mutation at the nucleotide level, the 2.8 kb *Pst*I

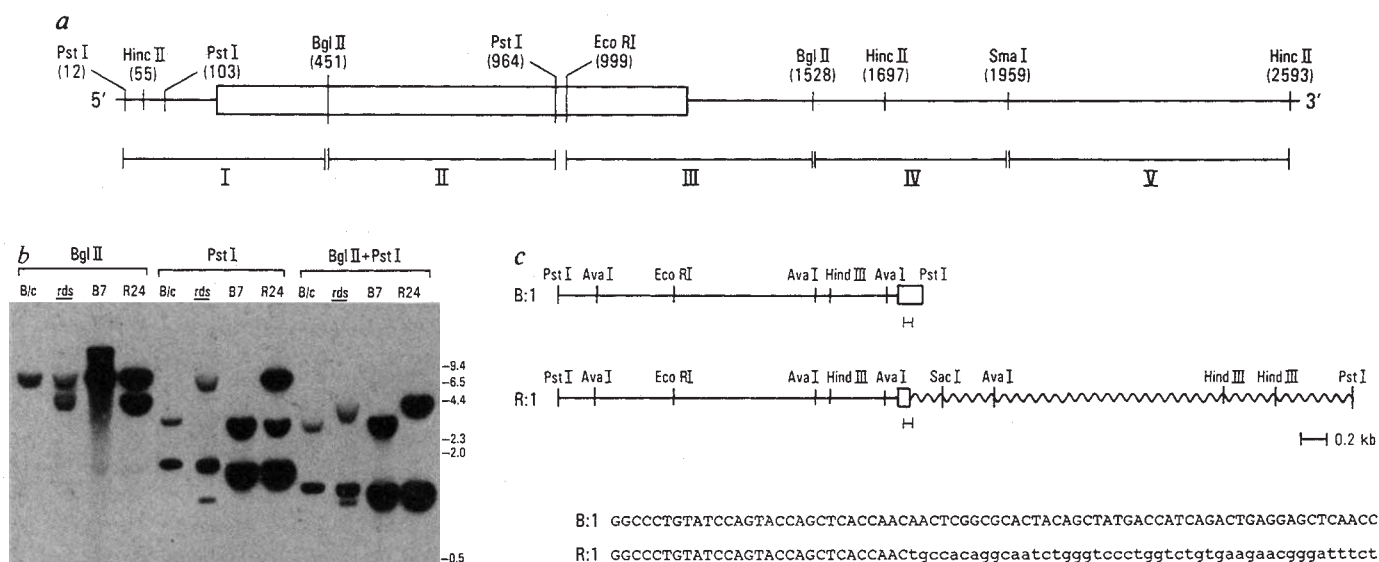


FIG. 2 *a*, Restriction map of the *rds* cDNA clone IG3.B9A derived from the nucleotide sequence (see Fig. 3). The open reading frame is indicated by the open box. Restriction fragments (B9A I-V) used as probes for Southern blot analysis are indicated below. *b*, Southern blot analysis of genomic and λ clone DNA with IG3.B9A fragment II probe. B/c, BALB/c DNA; *rds*, *rds/rds* DNA (10 μ g genomic DNA per lane); B7, λ clone B7 from BALB/c library; R24, λ clone R24 from *rds/rds* library (0.5 ng λ clone DNA plus 10 μ g sheared salmon sperm DNA per lane). DNA samples were digested with restriction endonuclease *Bgl*II, *Pst*I, or *Bgl*II plus *Pst*I as indicated. DNA size standards are indicated to the right in kb. Electrophoresis, transfer to nitrocellulose, preparation of probes and stringency wash conditions as described in Table 1.

c, Restriction map of BALB/c and *rds/rds* genomic clones. B:1 is the 2.8 kb *Pst*I genomic subclone fragment from wild-type phage λ clone B7 detected with probe B9A-II. R:1 is the 6.3 kb *Pst*I genomic subclone fragment from *rds/rds* phage λ clone R24 detected with probe B9A-II. The open box indicates an exon which contains a protein-coding region (see Fig. 3). The wavy line indicates the insertion in the mutant locus. The nucleotide sequences of B:1 and R:1 in the region containing the divergence is indicated below. Capital letters in the nucleotide sequence correspond to the wild-type exon; lower-case letters correspond to the insertion. Bars underneath the restriction map show the origin of the indicated sequences.

TABLE 1 Segregation pattern of the mouse allele

Chromosome	Hybridization pattern				Per cent discordance
	+/+	-/-	+/-	-/+	
1	4	5	4	1	35.7
2	5	3	3	3	42.9
3	2	4	3	0	33.3
4	3	3	5	2	53.8
5	0	6	8	0	57.1
6	3	4	5	2	50.0
7	4	0	4	6	71.4
8	3	5	5	0	38.5
9	1	5	7	1	57.1
10	1	6	7	0	50.0
11	0	4	6	0	60.0
12	4	1	2	3	50.0
13	3	2	5	3	61.5
14	1	5	7	1	57.1
15	6	0	0	4	40.0
16	2	4	5	1	50.0
17	8	6	0	0	0.0
18	3	2	4	3	58.3
19	2	2	6	3	69.2
X	4	5	4	1	35.7

Southern blots containing DNA from mouse $\times$ hamster somatic cell hybrids were probed with cDNA clone IG3. Column designations: +/+, chromosome and IG3 hybridization signal both present; -/-, chromosome and hybridization signal both absent; +/-, chromosome present and hybridization signal absent; -/+, chromosome absent and hybridization signal present. A hybridization signal corresponding to the hamster allele was present in all cell lines. 10 μ g DNA samples from 16 mouse $\times$ hamster hybrid cell lines containing different mouse chromosomes¹⁴ were digested with restriction endonuclease *Pst*I and resolved by electrophoresis on 0.7% agarose. The DNA was transferred to nitrocellulose¹⁷ and then probed with DNA from clone IG3 (see Fig. 1). Final stringency wash conditions were 0.2 $\times$ SSC/0.2% SDS at 68 $^{\circ}$ C.

fragment from (wild-type) λ clone B7 and the 6.3 kb *Pst*I fragment from (*rds/rds*) λ clone R24 detected by B9A-II were both subcloned. Restriction endonuclease mapping of these subcloned *Pst*I fragments (B:1 and R:1) revealed that the 5' 2.7 kb of each were indistinguishable. After the third *Ava*I site (Fig. 2c), the restriction maps of B:1 and R:1 diverged. The 0.25 kb *Ava*I-*Pst*I 3' fragment from the wild-type subclone B:1, and the corresponding 0.95 kb *Ava*I-*Ava*I fragment from the mutant subclone R:1 were isolated and their nucleotide sequences determined. The sequences were identical up to the position corresponding to nucleotide 899 in the mRNA (see below and Fig. 3). At this point, a foreign sequence appears in the mutant gene, disrupting the protein-coding exon (Fig. 2c). As the mutation occurs within an exon, this foreign sequence is probably transcribed in *rds/rds*. The observed increase in the size of the transcripts detected with clone IG3.B9A in the retinas of *rds/rds* mice is consistent with the insertion of approximately 10 kb of foreign DNA into an exon of the *rds* gene.

In an attempt to identify this foreign insert, we used Southern blots to analyse genomic DNA isolated from BALB/c and *rds/rds* using the 0.4 kb *Sac*I-*Ava*I restriction fragment from the inserted element of R:1 (Fig. 2c) as a probe. About 30 bands were detected in both the wild-type and *rds/rds* lanes, with two additional bands present in the *rds/rds* lanes that were absent from the wild-type lanes (data not shown). A computer database search was performed with the nucleotide sequence from the 0.4 kb *Sac*I-*Ava*I fragment of R:1 (data not shown). No significant similarity to any eukaryotic or viral sequences was detected. Based upon its size (approximately 10 kb) and its copy number in the mouse genome (10-30 copies), this inserted DNA may represent an unidentified retrovirus-like element.

Sequence analysis of the IG3.B9A cDNA revealed an insert containing 2,632 base pairs (bp) followed by a poly(A) tail (Fig. 3). A single long open reading frame began at nucleotide 213, encoding a putative protein of 346 amino acids. A common

FIG. 3 Nucleotide sequence of the full-length *rd*s cDNA clone IG3.B9A. Nucleotide numbering is shown on the left. The single long open reading frame has the corresponding amino-acid sequence translated above. The polyadenylation signal CATAAA for the 1.6 kb mRNA (IG3.D6), and ATTTAA for the 2.7 kb mRNA (IG3.B9A) are underlined. METHODS. The sequence was determined by a modified version²³ of the chemical degradation method of Maxam and Gilbert²⁴. The complete sequence was determined for both DNA strands of clones IG3.B9A and IG3.D6 (see below). Primer extension analysis was performed on clone IG3.B9A as described²⁵ using a synthetic deoxy-oligonucleotide corresponding to the antisense strand of IG3.B9A from position 64 to 41 (data not shown). cDNA clone D6 of the small, wild-type *rd*s transcript was isolated by screening for clones that hybridized with full-length clone IG3.B9A but failed to hybridize to a truncated clone containing 0.6 kb from the 3' end of IG3. The sequence of clone IG3.D6 was identical to IG3.B9A except for the use of an alternative polyadenylation signal.

```

AAGGACTCTCGAGATACGGCGGCTAGATTAGCTCCGGCTACCGTTACTGAGTTAACGGGGATCCCAAGCTAGGAGGCGCCCAAAATGG
90  GCAACTCCCTGCAGCTTGGGCGCATGCTGCTCTTCCCTAGACCTAGCGGTCCAGCCCGGAGGCTCACTCGGATTAGGATGGGAAGCTGA
MetAlaLeuLeuLysValLysPheAspGlnLysLysArgValLysLeuAlaGlnGly
180  ACCGTGGGAGGCTGTGAACGACCTCGGTAAAGCTGCTCAAGTCAAGTTTACACCAAGAAGCGGGTCAAGTTGGCCAGGCGG
LeuTrpLeuMetAsnTrpLeuSerValLeuAlaGlyIleValLeuPheSerLeuGlyLeuPheLeuGluLeuAlaLeuAsnTrpLysArgSer
270  CTCTGGCTTATGAACCTGGCTGCTGCTGCTGGCGGCTGCTCTCTCAGCTTGGGGCTGTTCTTGAAGATTGAATTCGCAAGAGGAGC
GluValMetAsnAsnSerGluSerHisPheValProAsnSerLeuIleGlyValGlyValLeuSerCysValPheAsnSerLeuAlaGly
360  GAAGTGATGAATAATCTGAGAGCCACTTGTGCGCACTCCCTGATAGGGGTGGGGGCTGCTGCTGTCTTCAACTCTCTGCTGCTGGG
LysIleCysTyrAspAlaLeuAspProAlaLysTyrAlaLysTrpLysProTrpLeuLysProTyrLeuAlaValCysIlePhePheAsn
450  AAGATCTGATGATGCCCTGGACCGGCAAGTACGCCAAGTGAAGCGTACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
ValIleLeuPheLeuValAlaLeuCysCysPheLeuLeuArgGlySerLeuGluSerThrLeuAlaTyrGlyLeuLysAsnGlyMetLys
540  GTCATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
TyrTyrArgAspThrAspThrProGlyArgCysPheMetLysLysThrIleAspMetLysGlnIleGluPheLysCysCysLysAsnAsn
630  TATTATCGGGATACGGACACCCCGGCGGCTGCTTATGAAAAAGACCATCCACATGCTCCAGATTGAGTCAAGTGTCTGGGAACAC
GlyPheArgAspTrpPheGluIleGlnTrpIleSerAsnArgTyrLeuAspPheSerSerLysLysValLysAspArgIleLysSerAsn
720  GGCTTCGGGAGCTGTTGAGATTACGATTCAGCAATCGCTACGCTGCTTCTCTCCAGGAGGCTCAAAAGTCCGATCCGATCAAGAGCAAT
ValAspGlyArgTyrLeuValAspGlyValProPheSerCysCysAsnProSerSerProArgProCysIleGlnTrpGlnLeuThrAsn
810  TGGATGGGCGGTACCTGGTGAAGCGGCTCCCTTTCAGCTGCTGCAACCCAGCTCCCGCGGCGCTGTATCCAGTACCAAGCTCACCAAC
AsnSerAlaHisTyrSerTyrAspHisGlnTrpGluGluLeuAsnLeuTrpLeuArgGlyLysAlaAlaLeuLeuAsnTrpLysSer
900  AACTCGGCGCACTACAGCTATGACCATCAGACTGAGGAGCTCAACCTCTGGTGGCGGCTGCAAGGCGCGCTGCTGCTGCTGCTGCTGCTG
SerLeuMetAsnSerMetGlyValValThrLeuLeuValTrpLeuPheGluValSerIleThrAlaGlyLeuArgTyrLeuHisThrAla
990  AGCTTCAGTAATTCATGGGCGCTGCACACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
LeuGluSerValSerAsnProGluAspProGluCysGluSerGluGlyTrpLeuGluGlySerValProGluTrpLysAlaPhe
1080  CTGGAGAGCTGTCTAACCCGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
LeuGluSerPheLysLysLeuGlyLysSerAsnGlnValGluAlaGluGlyAlaAspAlaGlyProAlaProGluAlaGly
1170  CTGGAGAGCTTTAAGAAGCTGGGCAAGCAATCAGGTGGAGGCTGAAGGTGCAGAGCGAGGCGCGGCTCCAGAGGCTGGCTGAGTGGCCT
GGGCGCTCTCCCTCTCTCAACACTTGTGAGCTCCAGGAGCTGTGAGTACCCCTTGTTCAGCTGAAAGTCCAAATTTCCGAGAAAG
1260  CTGCTCACTACTGACTCTCTTGTGAGTGGGCTTGAAGTTCAGGCTCCTTAGGCGAGCTATCAAACTTTGTGAAACGGCTGCTCCAG
1350  ATGTGATGATGTAAGATGACGAGTACGAGTGAACGCTGCAAGCTGCAAGGCTGCAAGGCTGCTGCAAGTGTGAGTGTGAGTGTG
1440  TCCTCATAGGTGACTGCCACACCAAGGGGCTCCCTCTCAGTATGTTGCTCTTTTAACTCAAGTTTGCATCCCAACCT
1530  CACTTGAACATATAAGCAAGGTGAAATAAGAAACTCTAAGGGCCATGTTGTTTCTTCTTAATGATAGGCTTAAAGCTGGTCA
1620  TCATTGCTCTCATGTGTACCAATGGGGGAGAGCAATGATATTTTAAAGATTTTGGACTCCCTCAGAAGTTACCTCCACATAGAAA
1710  TTCACAGATGGCAGAGGAGTAAATCTTCTTCTTATGAAATCCAGTGTGCTCCATCTCTGCGCCATCCAGAGCTCGGAGTCACT
1800  ACTTCAGCTGGCATGTCCAGCATCTCTCACTACCAAGTGTGGGTGAGCACTGCTGAGAAAGCGGCGGTGAGTTGCTGCTCTT
1890  GTGGGAGAGATGTGGTATCAATAGGAATCTGCACTTGGGGCAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1980  TGGATCTGCACTGTCCAGGAGGGCAAGCCCTGGAGCTTCTGTTCTCCCTCCCAAGCAAGCACTATGGGAATGCTTCCGTATCCAGGAA
2070  GATTCTGAGATAGTCTTCTTCCAGCTAGCTGTGTTGGAAGTTGCTGTTTCTTCTGGAACCTCAGGCAAACTCTCAGGGAAGCA
2160  AAAATGATTTTGTGGTGGCCTTGTGGGTGGCCACCCACAAGGACTGTTGTGGCTCATGAGACCCAGGAGCGGCTGTGCTGCTG
2250  AAGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2340  ATGCGCCCAATGTGGTATGATCTGATCTCTCAGAGCAAGAAATGGCCAGATGCTGAGCCAGGATCTGTGGCGGGGACCTCGGG
2430  GTGCTGCGCTCAGCTCTCAGGAGGGCTGAGACTATGTTGAAAGTACTCTAGACCCATTTCTAATGATCTGCTGCAACACTCGTATTAAGA
2520  GATATTAAGAACTATATTGTATG(A)n
2610

```

variant (AUUAAA)¹⁵ of the consensus polyadenylation signal (AAUAAA)¹⁶ was found 20 nucleotides upstream from the poly(A) tail. Primer extension analysis showed that IG3.B9A was only six nucleotides short of being a full-length clone of the 2.7 kb mRNA (see Fig. 3 legend). A cDNA clone of the 1.6 kb *rd*s transcript (IG3.D6) was isolated by further screening (Fig. 3 legend). Sequence analysis of IG3.D6 showed that the 1.6 kb mRNA was identical to the 2.7 kb mRNA except that the poly(A) tail began 24 bases downstream from the alternative polyadenylation signal CAUAAA¹⁵, at position 1,651 rather than 2,633.

A computer search of a protein sequence database (Protein Identification Resource, 31 March 1988) with the predicted protein encoded by the *rd*s mRNA revealed no significant similarities to any known protein sequences. The *rd*s protein (relative molecular mass, *M_r* 39,259) does not seem to have an N-terminal secretion signal sequence. It does contain three uncharged regions (amino acids 16–41, 100–122 and 252–275) which may represent membrane-spanning domains. Four sets of tandem basic residues (amino-acid residues: 11–13, 46–48, 178–179 and 324–325) are possible candidates for proteolytic cleavage signals. The protein also contains 13 cysteine residues which are potential sites for intra- or inter-chain disulphide linkages. As the *rd*s mRNA has been shown here to be photoreceptor-specific within the retina and is not detectable in brain, it is possible that the *rd*s protein function relates to some photoreceptor-specific process, possibly as an unidentified member of the visual transduction cascade. As the insertion occurred within a protein-coding exon, the resulting protein will be aberrant, lacking its normal C-terminal 87 amino-acid residues if translation of the aberrant *rd*s mRNA does occur in *rd*s/*rd*s retina.

From the evidence presented here, we concluded that we have cloned the *rd*s gene and characterized its normal mRNA products. This is the first molecular description of a gene determining neuronal degeneration in a mammalian system where nothing was known in advance about the gene product. It remains a

formal possibility that the gene identified here is not *rd*s: if loss of function of the identified gene product is not lethal to photoreceptors and if another gene on chromosome 17 which is also expressed in photoreceptors bears an independent second mutation. Final proof that the identified gene is indeed *rd*s will await correction of the phenotype by introducing DNA from the cloned wild-type allele into the *rd*s/*rd*s background in a transgenic experiment. Elucidation of the mechanism of photoreceptor degeneration in *rd*s/*rd*s will need studies of the *rd*s gene product. The first steps towards this involve immunocytochemical localization of the *rd*s protein within photoreceptors, and the identification of other proteins which may be physically associated. □

Received 7 October 1988; accepted 13 January 1989.

- Sanyal, S. & Jansen, H. *Neurosci Lett* **21**, 23–26 (1981).
- Cohen, A. I. *Invest. Ophthalmol. Vis. Sci.* **24**, 832–843 (1983).
- Jansen, H. G. & Sanyal, S. *J. comp. Neurol.* **224**, 71–84 (1984).
- Sanyal, S., DeRuiter, A. & Hawkins, R. K. *J. comp. Neurol.* **194**, 193–198 (1980).
- Van Nie, R., Ivanyi, D. & Demant, P. *Tissue Antigens* **12**, 106–108 (1978).
- Demant, P., Ivanyi, D. & Van Nie, R. *Tissue Antigens* **13**, 53–56 (1979).
- Sanyal, S. & Zeilmaker, G. H. *Expl Eye Res.* **39**, 231–246 (1984).
- Sanyal, S., Dees, C. & Zeilmaker, G. H. *J. Embryol. exp. Morph.* **98**, 111–121 (1986).
- Tansley, K. J. *Hered.* **45**, 123–127 (1954).
- Sidman, R. L. & Green, M. C. *J. Hered.* **56**, 23–29 (1965).
- LaVail, M. M. & Sidman, R. L. *Archs Ophthalmol.* **91**, 394–400 (1974).
- Carter-Dawson, L. D., LaVail, M. M. & Sidman, R. L. *Invest. Ophthalmol. Vis. Sci.* **17**, 489–498 (1978).
- Travis, G. H. & Sutcliffe, J. G. *Proc. natn. Acad. Sci. USA* **85**, 1696–1700 (1988).
- Hoggan, M. D., Halden, N. F., Buckler, C. E. & Kozak, C. A. *J. Virol.* **62**, 1055–1056 (1988).
- Birnstein, M. L., Busslinger, M. & Strub, K. *Cell* **41**, 349–359 (1985).
- Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211–214 (1976).
- Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
- Danielson, P. E. et al. *Proc. natn. Acad. Sci. USA* **75**, 261–267 (1978).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
- Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. USA* **69**, 1408–1412 (1972).
- Thomas, P. S. *Proc. natn. Acad. Sci. USA* **77**, 5201–5205 (1980).
- Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6–13 (1983).
- Brow, M. A. D., Pesin, R. & Sutcliffe, J. G. *Molec. Biol. Evol.* **2**, 1–12 (1985).
- Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. USA* **74**, 560–564 (1977).
- Giorgi, C., Blumberg, B. & Kolakofsky, D. *Cell* **35**, 829–836 (1983).

ACKNOWLEDGEMENTS. We thank Sonja Forss-Petter for assistance in formulating the sequencing strategies for the cDNA clones, Adam Travis for useful comments on the manuscript and Somers Sanyal (supported by the NEI) for providing breeding pairs of O20/cpb (*rd*s/*rd*s) and BALB/c background strain mice. This work was supported by the NIH.

Does T-cell tolerance require a dedicated antigen-presenting cell?

Polly Matzinger & Sylvie Guerder

Basel Institute for Immunology, Grenzacherstrasse 487,
CH-4005 Basel, Switzerland

ALMOST 30 years ago Burnet proposed that the immune system maintained self-tolerance by deleting autoreactive lymphocytes¹. Recently it has become clear that for T cells this step occurs in the thymus, where developing T cells first express their antigen-specific receptors²⁻⁴. Here a T-cell which encounters its antigen disappears—if it is not dead, it at least stops expressing its receptors. In the periphery by contrast, encounter with antigen leads to activation and proliferation of the responding T-cell. There are two possible explanations for this difference. Either the antigen-presenting cells in the thymus are different from those in the periphery and instead of producing positive signals they directly or indirectly kill the thymocytes^{5,6}; or the T cells themselves are different, and like immature B cells, may die after encounter with antigen^{7,8}. We tested the first possibility and found that dendritic cells from spleen, which are the most potent activators of mature T cells⁹, are also the most potent inactivators of young developing T cells. Thus it is not the antigen-presenting cell which determines whether a T-cell responds or dies, but the T-cell itself or its thymic environment.

To analyse the activity of various cell types, we used *in vitro* thymus organ cultures^{10,11}. We cultured embryonic mouse thymus lobes from 14-day-old fetal B6 or (B6 × BALB/c) mice for 7 days. At day 14 of gestation these thymuses have no mature T cells^{12,13}, but by day 21 the T cells have matured¹¹ and can be activated by mitogens¹¹ or in mixed-lymphocyte reactions¹⁴.

Figure 1 shows the result of two experiments designed to find out whether it was possible to tolerize cytotoxic T cells developing *in vitro*. Since it had been shown that proliferating T cells can be rendered specifically unresponsive if they are co-cultured with allogeneic adult¹¹ or fetal¹⁵ thymus cells, we cultured the fetal thymuses in contact with a slice of irradiated adult allogeneic (SJL or CBA) thymus. Figure 1a and c shows that fetal T cells can develop into cytotoxic T lymphocytes (CTL) specific for allogeneic C3H.Q, CBA or SJL stimulators. Fig. 1b and d demonstrates that co-culture with adult thymic tissue generates specific unresponsiveness. Thymuses co-cultured with CBA (Fig. 1b) no longer generated CTL against CBA, but remained totally responsive to C3H.Q, whereas thymuses co-cultured with SJL became tolerant of SJL while maintaining their activity against CBA (Fig. 1d). Using limiting-dilution analysis (not shown), we found that the frequency of cells responding to the co-cultured thymus was decreased 100- to 1000-fold. Thus it was clear that the *in vitro* development of immature T cells is sufficiently similar to their *in vivo* development to allow both maturation and tolerance of specific CTL.

Although it has been assumed for some time that the 'tolerizing' antigen-presenting cell (APC) in the thymus is a bone marrow-derived dendritic cell¹⁵, a VETO cell⁵ or a double-

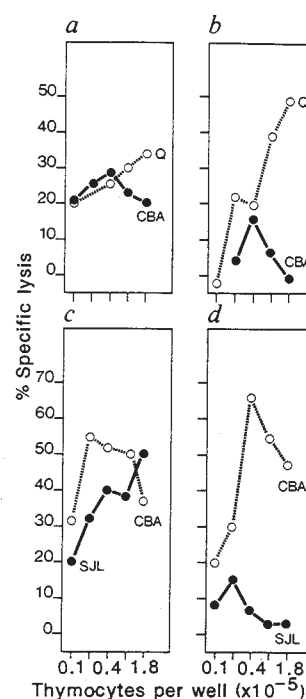


FIG. 1 Tolerance *in vitro*. Details of the mouse strains are shown in Table 1. We removed thymuses from 14-day (B6 × BALB/c) mouse fetuses and cultured them on millipore filters at the liquid-air interface in Iscove's modified Dulbecco's medium with 10% fetal calf serum, antibiotics and 6×10^{-5} M α -thioglycerol¹¹. The thymuses were either cultured alone (a, c; 8–10 per filter) or around and in contact with a slice (~1 mm thick) of irradiated (1,000 rads) adult CBA (b) or SJL (d) thymus. Activation cultures and cytotoxic assay: After 7 days thymocytes were collected and graded (indicated as numbers of thymocytes per well) into round-bottomed 96-well microtitre plates with 5×10^5 irradiated (3,000 rad) adult spleen stimulator cells in medium including 10% supernatant from concanavalin-A activated rat spleen cells plus 20 mM α -methyl mannose. After 6 days the medium was removed and replaced with 0.2 ml medium containing 10^4 ^{51}Cr -labelled targets (spleen cells cultured for 2 d with $2 \mu\text{g ml}^{-1}$ concanavalin A). Released ^{51}Cr was collected after 3.5 h. Cytotoxicity (per cent) was averaged from triplicate cultures and calculated from: (experimental release – spontaneous release)/(maximum release – spontaneous release) $\times 100$. Lytic activity is shown against the stimulating cell type. The activity on inappropriate targets was never >20% of the specific activity and was usually <1%. Occasionally (~1/8 experiments) the level of nonspecific cytotoxicity is anomalously high and so these experiments are excluded. The somewhat unusual shape of the cytotoxicity curves probably results from two factors: at low cell numbers we are at limiting dilution conditions and single active clones can sporadically appear to generate unusual peaks of activity; also, in a very active response, high cell numbers will deplete the culture medium and die before the assay. We would therefore see efficient CTL activity only at lower inputs. a, c, Thymuses cultured alone; b, co-cultured with CBA; d, co-cultured with SJL.

negative thymocyte¹⁶, it is not known whether the tolerizing cell must come from the thymus or whether the newly developing T-cell can be rendered unresponsive after interaction with any APC. To distinguish between the two alternatives, we tested the tolerizing capacity of dendritic cells from adult spleen. In the activation of mature T cells, splenic dendritic cells are among the most efficient of APCs⁹. Would they activate or tolerize newly developing thymocytes?

Figure 2a shows the responses of (B6 × BALB/c) thymocytes which have developed alone or in the presence of graded numbers of irradiated dendritic cells from the spleens of adult CBA mice. Thymocytes that developed alone generated good CTL against both CBA and C3H.Q, whereas thymocytes that developed in the presence of even small numbers of CBA dendritic cells responded to C3H.Q but were no longer respon-

TABLE 1 Details of mouse strains

Strain	Abbreviation	H-2 type
C57Bl/6	B6	b
BALB/c	C	d
C57Bl/6 × BALB/c	(B6 × C)	b × d
CBA	CBA	k
C3H	C3H	k
C3H.Q	Q	q
SJL	SJL	s

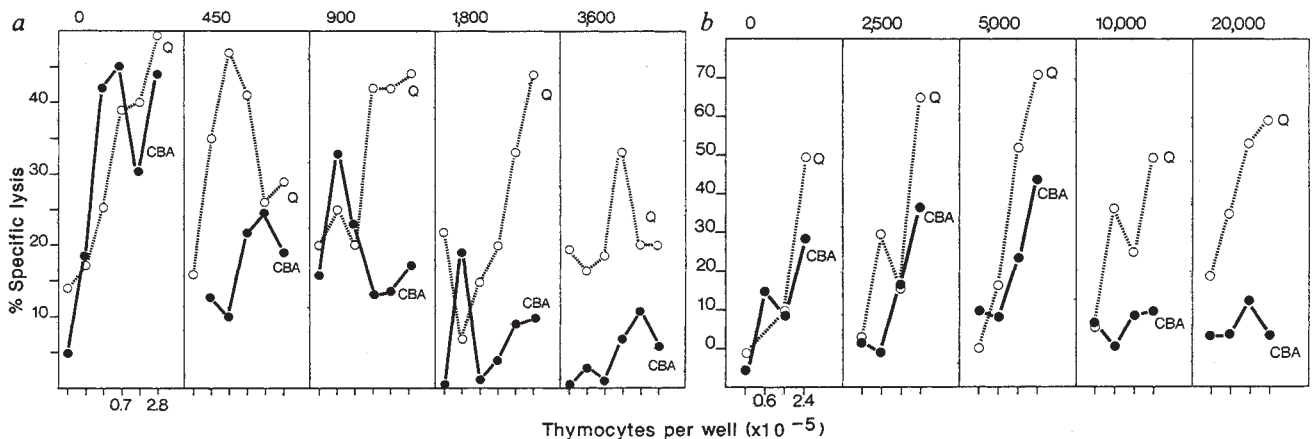


FIG. 2 Dendritic cells induce tolerance *in vitro* and *in vivo*. *a*, Thymic organ culture *in vitro*. Each 14-day fetal thymus lobe was cultured in a 20 μ l Terasaki well (upside down) in the presence or absence of graded numbers of irradiated (1,000 rads) CBA dendritic cells (20 thymus lobes per group). After 18 h each lobe was rinsed, transferred to a Millipore filter, and cultured for another 6 d. For the activation culture the 20 lobes in each group were pooled, activated and tested as described in Fig. 1 legend. The number of dendritic cells per lobe is given above each panel. *b*, Thymic organ culture *in vivo*. Thymuses were cultured overnight with CBA dendritic cells. They were then rinsed and grafted under the left kidney capsule of adult SCID mice

(8–10 lobes per mouse, 2 mice per group). After 11 days the thymocytes were collected, pooled, activated and tested. Isolation of dendritic cells: $1-2 \times 10^8$ spleen cells per dish were cultured in 5 ml Petri dishes for 2 h, stringently washed free of all non-adherent cells. The remaining cells were then re-cultured for 18 h and the non-adherent cells gently collected. The yield varied from 0.1–0.3% of input, contained no cells positive for Thyl or immunoglobulin and <0.01% Fc receptor-positive cells. In our hands these dendritic cells were 20–30 fold more effective than spleen cells at presenting antigen to T-cell clones.

sive to CBA. Figure 2b shows that dendritic cells have the same effect *in vivo*. We added graded numbers of CBA dendritic cells, cultured the thymuses overnight and then allowed them to develop for 11 days under the kidney capsules of SCID mice (which cannot generate their own T or B cells¹⁷). Again, the presence of adult spleen dendritic cells specifically abolished responsiveness.

The number of dendritic cells which induce unresponsiveness is small and is correlated with the size of the thymus. *In vitro* the thymuses develop $2-5 \times 10^5$ thymocytes per lobe, whereas *in vivo* the same thymuses will enlarge to $1-4 \times 10^6$. This 10-fold difference is reflected in the number of allogeneic dendritic cells needed for tolerance. On the average, complete tolerance is achieved with one dendritic cell per 200 thymocytes *in vitro* and *in vivo*. In tissue sections (not shown) we found that, like dendritic cells in normal thymus¹⁸, our dendritic cells do not distribute randomly throughout the thymus, but seem to collect into tubules or clusters in the medulla and at the cortico-medullary junction. Perhaps young T cells must run a gauntlet of dendritic cells before they are allowed to leave the thymus.

Figure 3 indicates that other cell types are not as effective as dendritic cells. It takes $1-2 \times 10^3$ dendritic cells to remove 90% of the thymocyte responsiveness against them, but 100-fold more spleen cells and 1,000-fold more thymocytes are required to achieve the same effect. This difference correlates well with the frequency of dendritic cells in each population⁹, implying that antigen presentation for tolerance, like antigen presentation for activation, is best done by a typical APC, and that dendritic cells, the most effective activators of mature T cells, are also the most effective inactivators of young thymocytes.

The tolerizing ability of dendritic cells could solve two long standing problems. The first involves tolerance to peripheral tissue-specific antigens which are not expressed by the thymus^{19,20}. It cannot be argued that the thymus specializes in the expression of each possible self-antigen, because tetraparental embryo-fusion chimaeras express antigens on their syncytiated muscle cells that cannot be expressed elsewhere (including in the thymus)¹⁹, and yet these mice are tolerant of their muscle. Since peripheral APCs circulate through the thymus²¹, they could well pick up a variety of antigens on the way, which they could then present to the developing thymocytes, thereby generating at least some level of tolerance to peripheral antigens.

Of course this does not exclude the existence of peripheral suppressive 'fail-safe' mechanisms to ensure self-tolerance should thymic depletion be imperfect³². The second problem concerns extrathymic differentiation²². Nude mice, for example, have no thymuses, yet small numbers of T cells can be found. These T cells can be activated against allogeneic and haptenated targets²³ but they are totally tolerant of self. If the decision to respond or be turned off is actually made by the developing T-cell rather than by the APC, then any T cells developing outside the thymus will also be rendered self-tolerant. Thus the system provides a safety valve that allows for extrathymic differentiation of T cells (and for potential evolutionary side-steps).

Tolerance induction by normal peripheral APCs also allows for conservation of specificity. We know that tolerance, like

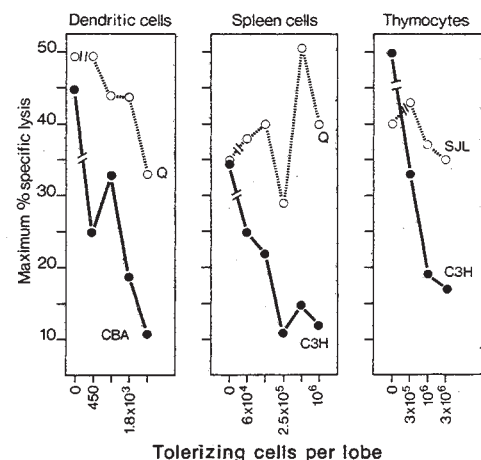


FIG. 3 Comparison of the effect of different cell populations. 14-day (B6 $\times$ BALB/c) thymuses were cultured with graded numbers of irradiated adult CBA spleen or thymus cells or purified dendritic cells as described in Fig. 1 legend. The ordinate represents the highest level of cytotoxicity seen in the titrated culture of each (pooled) group of thymocytes; the abscissa represents the number of 'tolerizing' cells per lobe. For comparison, the first panel is the same experiment as that shown in Fig. 2a. The highest point in each panel of Fig. 2a becomes one point on the curve here.

activation, is MHC-restricted^{24,25}, and that antigen-processing is probably involved^{26,27}. According to Bretscher and Cohn²⁸ 'every inducible cell must be tolerizable' and it should be tolerizable by the same antigen, at the same concentration and presented in the same way as the one to which it could be activated. If T cells are normally activated by peripheral APCs then it makes sense to tolerize them using the same APCs.

Here we have established that tolerance is not uniquely the result of antigen presentation by a special set of thymic antigen-presenting cells. Young thymocytes can be turned off when they recognize antigen on the same types of APCs that activate mature T cells. It could be argued that the peripheral dendritic cell might change its role on entering the thymus, turning from an activator into a tolerizer. Two sets of results stand against this. First, APCs isolated from the thymus can stimulate T-cell clones²⁹. Second, mature T cells in the thymus can be primed by a peripheral injection of spleen cells³⁰. We favour the view that the decision for tolerance, as opposed to activation, is made by the T cells themselves at different stages of their development. Whether the young thymocytes are innately suicidal or in need of critical, but unavailable, growth factors³¹ remains to be investigated. □

Received 25 July 1988; accepted 3 January 1989.

1. Burnet, F. M. *The Clonal Selection of Acquired Immunity* (Cambridge University Press, New York, 1959).

2. Kappler, J. W. *et al. Nature* **332**, 35–40 (1988).
3. MacDonald, H. R. *et al. Nature* **332**, 40–45 (1988).
4. Kisielow, P. *et al. Nature* **333**, 742–746 (1988).
5. Miller, R. G. *Nature* **287**, 544–546 (1980).
6. Rammensee, H.-G., Bevan, M. J. & Fink, P. *J. Immun. Today* **6**, 41–43 (1985).
7. Nossal, G. J. V. & Pike, B. *J. exp. Med.* **141**, 904–917 (1975).
8. Metcalf, E. S. & Kinman, N. R. *J. exp. Med.* **143**, 1327–1340 (1976).
9. Steinman, R. M. & Witmer, M. D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5132–5136 (1978).
10. Robinson, J. H. & Owen, J. J. T. *Clin. exp. Immun.* **27**, 322–327 (1977).
11. Robinson, J. H. & Owen, J. J. T. *Nature* **271**, 758–760 (1978).
12. Ceredig, R. *et al. J. exp. Med.* **158**, 1654–1671 (1983).
13. Snodgrass, R. *et al. Nature* **315**, 232–233 (1985).
14. Kisielow, P., Leiserson, W. & von Boehmer, H. *J. Immun.* **133**, 1117–1123 (1984).
15. Jenkinson, E. J., Jhittay, P., Kingston, R. & Owen, J. J. T. *Transplantation* **39**, 331–333 (1985).
16. Shimorkevitz, R. P. & Bevan, M. J. *J. exp. Med.* **168**, 143–156 (1988).
17. Bosma, C. G., Custer, R. P. & Bosma, M. J. *Nature* **301**, 527–528 (1983).
18. Barclay, N. & Maryhoffer, G. *J. exp. Med.* **153**, 1666–1671 (1981).
19. Minz, B. & Baker, W. W. *Proc. natn. Acad. Sci. U.S.A.* **58**, 592–598 (1967).
20. Boyse, E. A. *et al. Nature* **227**, 901–903 (1970).
21. Longo, D. L. & Schwartz, R. H. *Nature* **287**, 44–46 (1980).
22. Kruisbeek, A. M. *et al. J. exp. Med.* **160**, 839–857 (1984).
23. Hunig, T. & Bevan, M. J. *Exp. Cell Biol.* **52**, 7–11 (1984).
24. Groves, E. & Singer, A. *J. exp. Med.* **158**, 1483–1497 (1983).
25. Matzinger, P., Zamoyka, R. & Waldmann, H. *Nature* **308**, 738–741 (1984).
26. Matzinger, P. & Waterfield, J. D. *Nature* **285**, 492–494 (1980).
27. von Boehmer, H. & Hafen, K. *Nature* **320**, 626–628 (1986).
28. Bretscher, P. & Cohn, M. *Science* **169**, 1042–1049 (1970).
29. Longo, D. L. & Schwartz, R. H. *Nature* **287**, 44–46 (1980).
30. Fink, P. J., Bevan, M. J. & Weissman, I. L. *J. exp. Med.* **159**, 36–41 (1984).
31. Malkovsky, M. & Medawar, P. B. *Immun. Today* **5**, 340–343 (1984).
32. Zamoyka, R., Waldmann, H. & Matzinger, P. *Eur. J. Immun.* (in the press).

ACKNOWLEDGEMENTS. We thank our colleagues David Gray, James Kaufman, Charles W. B. Mackay, Fritz Melchers, Ronald Palacios, Charley Steinberg and Harald von Boehmer for suggestions, criticisms and support, and Janette Millar for typing the manuscript. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche & Co. Ltd, Basel, Switzerland.

Positive selection of CD4⁺CD8⁺ T cells in the thymus of normal mice

Juan Carlos Zúñiga-Pflucker, Dan L. Longo & Ada M. Kruisbeek*

Biological Response Modifiers Program, National Cancer Institute, Building 10, Room 12N226, National Institutes of Health, Bethesda, Maryland 20892, USA

THE diversification of the repertoire of T-cell antigen receptor (TCR) specificities is influenced by at least two selection processes which occur in the thymus^{1–6}. One of these, termed 'negative selection', is required to install a state of tolerance to self-antigens in the T-cell repertoire and is often achieved by clonal deletion^{7–11}. The second type of selection operating in the thymus results in preferential differentiation of T cells that have restriction specificity for thymic major histocompatibility complex glycoproteins, but the mechanisms leading to this selective process are not yet clear. One model used to describe this 'positive selection' proposes that only those T cells with sufficient avidity for the MHC glycoproteins expressed in the thymus are allowed to acquire functional competence^{1–4,12,13}. Here we directly investigate the generation of TCR specificities by following the fate of developing Vβ17⁺ CD4⁺CD8⁺ T cells under conditions where one of the main class I-MHC molecules, either H-2K or H-2D, was specifically blocked by *in vitro* monoclonal antibody treatment¹⁴. The results show that development of Vβ17⁺ CD4⁺CD8⁺ T cells in the SJL H-2^s mouse strain is selectively abrogated by blocking class I-K^s molecules but is unaffected by blocking class I-D^s molecules. These data directly demonstrate that generation of CD4⁺CD8⁺ T cells expressing a particular TCR Vβ segment can be correlated with the expression of a particular class I-MHC molecule, thereby providing evidence for positive selection.

The TCR-Vβ17a product confers reactivity to various allelic forms of the murine class II-MHC molecule I-E (refs 7 and 15)

and is differentially expressed among the CD4⁺CD8⁺ and CD4⁺CD8⁺ T-cell subsets of different mouse strains. For instance, SWR (H-2^a) mice express Vβ17a predominantly in CD4⁺CD8⁺ T cells, whereas SJL (H-2^s) mice express Vβ17a in both CD4⁺CD8⁺ and CD4⁺CD8⁺ T cells^{7,16}. One possible explanation for this is that the class I-MHC glycoproteins of the H-2^s haplotype, but not those of the H-2^a haplotype, allow for positive selection of the Vβ17⁺ CD4⁺CD8⁺ T cells. An inference from this hypothesis is that blocking one of the H-2^s-class I MHC molecules could prevent the generation of Vβ17⁺ CD4⁺CD8⁺ T cells in H-2^s mice without affecting the overall development of CD4⁺CD8⁺ T cells.

To test this prediction, we treated SJL (H-2^s) mice from the day of birth with anti-class I monoclonal antibody (mAb) against either K^s (mAb 34-1-2; ref. 17; mouse γ2a) or D^s (mAb 15-1-5; ref. 18; mouse γ2b). The doses used (500 μg per gram body weight administered daily until analysis at 2–3 weeks of age)¹⁴ were sufficient to saturate the appropriate class I-MHC on spleen

TABLE 1 Expression of Vβ17 by thymocytes from control, anti-K^s-treated, and anti-D^s treated H-2^s mice

Treatment	Vβ17a in CD3 ⁺ CD4 ⁺ CD8 ⁺ cells	Vβ17a in CD3 ⁺ CD4 ⁺ CD8 ⁺ cells
Experiment 1		
Control	5.7	23.5
Anti-K ^s	0.3	23.9
Anti-D ^s	5.9	22.6
Experiment 2		
Control	5.4	20.9
Anti-K ^s	0.1	22.3
Anti-D ^s	4.8	22.4
Experiment 3		
Control	4.8	19.2
Anti-K ^s	0.1	23.2
Anti-D ^s	4.6	20.6

Mice were treated with anti-K^s or anti-D^s mAb as for Fig. 1. Thymocytes were enriched for receptor-bearing single positive cells by treatment with M1/69 plus complement as described in Fig. 2. Data are presented as per cent positive cells (values for control-staining subtracted).

*To whom correspondence should be addressed.

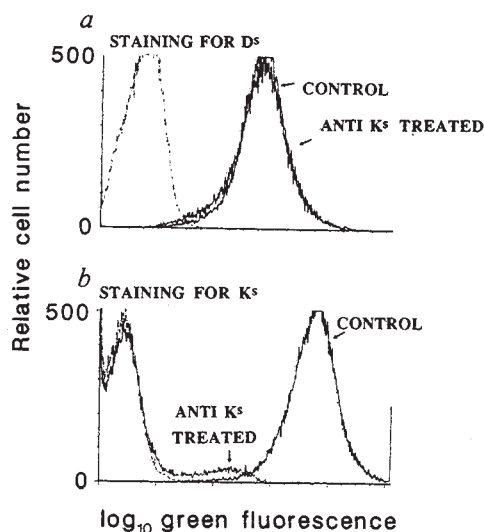


FIG. 1 Flow cytometry analysis of cell-surface class I-H-2D^s and class I-H-2K^s expression on thymocytes from control and anti-K^s-treated mice. Cells were stained with anti-D^s-mAb (a, solid lines) or anti-K^s-mAb (b, solid lines). Staining with irrelevant control mAb (dashed lines) is shown for only one of the cell samples (control or anti-K^s-treated).

METHODS. Neonatal SJL (H-2^s) mice were treated within 24 h of birth with anti-K^s mAb (hybridoma 34-1-2, ref. 17; reactive with K^sD^s and cross-reactive with K^s) or anti-H-2D^s mAb (hybridoma 15-1-5, ref. 18; reactive with K^sD^s and cross-reactive with D^s). Hybridomas were obtained from ATCC (American Type Culture Collection) and purified from ascites by ammonium sulphate precipitation followed by size separation. Mice were injected daily with 500 µg mAb per gram body weight and their thymocytes analysed at 2–3 weeks of age in 4 separate experiments. Thymocytes were pretreated *in vitro* with mAb M1/69 (ref. 20) plus complement to enrich for mature T cells (see Fig. 2), and stained with either fluorescein isothiocyanate-labelled 15-1-5 (anti-D^s) mAb (a) or biotinylated 34-1-2 (anti-K^s) mAb, followed by allophycocyanin-labelled avidin (Caltag, San Francisco) (b). Flow cytometry analysis was performed on a B-D Dual Laser 440 interfaced to a PDP 11/24 computer¹⁴. Data were collected on 50,000 viable cells and are shown for control and anti-K^s-treated mice. Reciprocal results are obtained in anti-D^s treated mice, in that D^s-staining is reduced and K^s-staining is normal (data not shown). Staining with appropriate anti-mouse IgG antibody revealed that reduced staining was due to blocking by mAb (not shown). Control mice were treated with saline, once preliminary experiments had shown¹⁴ that daily injection with equally high amounts of non-binding anti-class I mAb had no effect

cells (data not shown) and thymocytes (Fig. 1). Flow cytometry analysis of anti-K^s-treated mice revealed there had been a strong blocking of K^s while D^s-expression was left undisturbed. Reciprocal results, with D^s-blocking and normal K^s expression, were observed in anti-D^s-treated mice (data not shown).

We next investigated Vβ17a expression in developing thymocytes from H-2^s mice treated from birth with anti-K^s or anti-D^s. Although treatment of newborn homozygous mice with anti-class I mAb specific for both H-2K and H-2D-encoded glycoproteins prevents generation of most CD4⁺CD8⁺ T cells¹⁴, treatment of newborn F₁ mice with anti-class I mAbs that are specific for only one of the two parental haplotypes has no effect on the total number of CD4⁺CD8⁺ T cells¹⁴. By analogy, treatment of homozygous mice with anti-class I mAb specific only for either H-2K- or H-2D-region-encoded products has no effect on overall development of mature CD3⁺CD4⁺CD8⁺ T cells¹⁹ (see also Fig. 2). Thus, selective anti-K^s or anti-D^s mAb treatment of H-2^s mice will allow for TCR-repertoire analysis of developing CD4⁺CD8⁺ T cells. To enrich for TCR-bearing cells, thymocytes from control or treated mice were subjected to M1/69 (ref. 20) plus complement *in vitro* before analysis of TCR expression. The M1/69 mAb recognizes an epitope on the same molecule as the JIId mAb (ref. 21) and eliminates almost all CD4⁺CD8⁺ thymocytes (see also Fig. 2) and a subset of CD4⁺CD8⁺ thymocytes²². Anti-D^s-mAb treatment has no effect

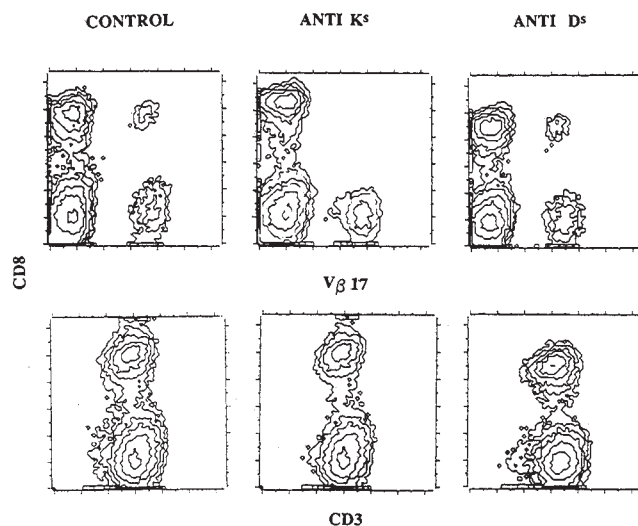


FIG. 2 Two-colour flow cytometry analysis of cell-surface CD8 expression versus Vβ17a-TCR (top) or CD3 (bottom) on thymocytes from SJL (H-2^s) mice which were untreated (left), treated from birth with anti-K^s mAb (middle), or treated from birth with anti-D^s mAb (right). Cells were stained first with FITC-conjugated anti-Vβ17 mAb or FITC-conjugated anti-CD3 mAb, then washed and stained with biotinylated anti-CD8 mAb followed by allophycocyanin-labelled avidin.

METHODS. Mice were treated with anti-K^s or anti-D^s mAbs until they were 3 weeks old, as described in Fig. 1. Thymocytes were treated with mAb M1/69 (ref. 20) plus complement to enrich for mature T cells, and stained for Vβ17a-expression with FITC-conjugated KJ23 (ref. 15) or for CD3-expression with FITC-conjugated 145-2C11 (ref. 32). Cells were then washed and stained for CD8 with biotinylated 53.6 mAb. Data were collected as described in Fig. 1 and displayed as contour diagrams. Per cent Vβ17a-expressing cells: for CD3⁺CD4⁺CD8⁺ cells, 5.7% in control; <0.5% in anti-K^s-treated; 5.9% in anti-D^s-treated. Per cent Vβ17a-expressing cells in CD3⁺CD4⁺CD8⁺ cells: 23.5% in control; 23.9% in anti-K^s-treated; 22.6% in anti-D^s-treated; these last values were derived from Vβ17a versus CD4 staining (not shown).

on Vβ17a-expression in either one of the subsets of mature T cells (Fig. 2, top). In striking contrast, thymocytes from anti-K^s treated mice exhibited virtually no Vβ17a-expression in the CD4⁺CD8⁺ subset (Fig. 2, top) (<0.5% in treated versus 5.7% in control); a summary of three experiments is shown in Table 1. The effect of the anti-K^s treatment on Vβ17a-TCR bearing CD4⁺CD8⁺ cells was specific in that it did not affect the total number of CD3⁺CD4⁺CD8⁺ T cells (Fig. 2, bottom), nor did it have any effect on the generation of Vβ17a⁺CD4⁺CD8⁺ cells (Fig. 2, top; 23.9% in treated versus 23.5% in control). This highly selective perturbation of the TCR-repertoire of CD4⁺CD8⁺ T cells is strong evidence that development of CD4⁺CD8⁺ T cells with a certain TCR requires expression of particular alleles of the class I-MHC glycoproteins. These findings indicate that TCR-MHC ligand interactions provide signals essential to the differentiation of precursor T cells, and are evidence in favour of the positive-selection theory.

Our results on T-cell repertoire development in anti-class I-treated normal mice confirm the proposed thymic selection inferred indirectly from chimera models¹⁻⁴. The data are also consistent with the observation that anti-class II-treated F₁ mice, in which the mAb can react with only one of the parental haplotypes, express a significantly altered repertoire²³, with T-helper cells showing self-class II restriction for the unblocked class II-molecules only. In a third model system using αβ-TCR-transgenic mice^{10,24-26}, positive selection in the context of thymic MHC also occurs, a result which together with ours shows the importance of TCR-MHC interactions in the development of T cells. Our data also bear on another aspect of TCR-repertoire selection, namely the degree to which certain TCR V gene elements can independently account for preferential recognition

of particular alleles of MHC glycoproteins. Several of the murine V β genes so far examined have, when expressed on CD4⁺CD8⁺ T cells, been associated with recognition of particular class II MHC molecules and self-antigens^{7-9,27-29}. The observation that V β 17a-bearing CD4⁺CD8⁺ T cells are not generated in anti-K^s-treated H-2^s mice leads to the prediction that V β 17a alone might confer restriction specificity for K^s, regardless of the V α sequences expressed on its partner chain and of junctional sequences. Thus, all V β 17a⁺ CD4⁺CD8⁺ T cells are expected to be K^s-restricted. Although this prediction can be tested by analysing the restriction specificities of various antigen-specific V β 17a-bearing CD4⁺CD8⁺ T cells derived from H-2^s mice, we have not yet been able to generate such cells, presumably because the vast number of potential foreign antigens precludes an easy definition of the appropriate ones. Nonetheless, the most straightforward explanation of our results is that V β 17a⁺ CD4⁺CD8⁺ T cells are preferentially selected on K^s, or by K^s complexed to self-peptides^{30,31}. Another explanation arises if we consider the opposite possibility that expression of certain TCR V gene elements precludes recognition of particular MHC molecules. According to such a model, all V β 17a CD4⁺CD8⁺ are apparently selected on K^s not because V β 17a 'dictates' K^s recognition, but because a TCR with V β 17a cannot interact with D^s; complexing of D^s with those self-peptides allowing for interactions with V β 17a (refs 30 and 31) could be inefficient. We consider this hypothesis attractive because other class I-gene products seem to be 'nonpermissive' for interactions with V β 17a: H-2^a mice express V β 17a predominantly in CD4⁺CD8⁺ cells, but do not express self-genes leading to V β 17a⁺-deletion in the CD4⁺CD8⁺ population (that is, (SWR $\times$ SJL) F₁ mice (H-2^a $\times$ H-2^s) do not delete V β 17a⁺ CD4⁺CD8⁺ T cells; unpublished observations). This is most easily explained by postulating that K^a and D^a do not allow for interactions with V β 17a, so V β 17a CD4⁺CD8⁺ T cells cannot be positively selected. We therefore suggest that the intrinsic lack of affinity of a certain TCR V β gene for particular MHC molecules may sometimes 'skew' the developing TCR repertoire and generate an apparent rather than a real preference for recognition of other MHC glycoproteins. □

A membrane form of guanylate cyclase is an atrial natriuretic peptide receptor

Michael Chinkers*†, David L. Garbers*†‡, Ming-Shi Chang§, David G. Lowe§, Hemin Chin||, David V. Goeddel§ & Stephanie Schulz*†

* The Howard Hughes Medical Institute, † Departments of Pharmacology and ‡ Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA

§ Department of Molecular Biology, Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

|| Laboratory of Biochemical Genetics, National Institutes of Health, Bethesda, Maryland 20892, USA

ATRIAL natriuretic peptide (ANP) is a polypeptide hormone whose effects include the induction of diuresis, natriuresis and vasorelaxation¹. One of the earliest events following binding of ANP to receptors on target cells is an increase in cyclic GMP concentration, indicating that this nucleotide might act as a mediator of the physiological effects of the hormone^{2,3}. Guanylate cyclase exists in at least two different molecular forms: a soluble haem-containing enzyme consisting of two subunits^{4,5} and a non-haem-containing transmembrane protein having a single subunit⁶. It is the membrane form of guanylate cyclase that is activated following binding of ANP to target cells^{3,7,8}. We report here the isolation, sequence and expression of a complementary DNA clone encoding the membrane form of guanylate cyclase from rat brain. Transfection of this cDNA into cultured mammalian cells results in expression of guanylate cyclase activity and ANP-binding activity. The ANP receptor/guanylate cyclase represents a new class of mammalian cell-surface receptors which contain an extracellular ligand-binding domain and an intracellular guanylate cyclase catalytic domain.

Using as probe a cDNA encoding the membrane form of guanylate cyclase from the sea urchin *Arbacia punctulata*⁶, we isolated partial-length putative guanylate cyclase clones from

Received 19 September 1988; accepted 11 January 1989.

- Bevan, M. J. *Nature* **269**, 417-419 (1977).
- Fink, P. & Bevan, M. J. *Exp. Med.* **148**, 766-775 (1978).
- Zinkernagel, R. M., Callahan, G. N., Klein, J. & Dennert, G. *Nature* **271**, 251-253 (1978).
- Zinkernagel, R. M. *et al.* *J. exp. Med.* **147**, 882-896 (1978).
- Kruisbeek, A. M., Hodes, R. J. & Singer, A. *J. exp. Med.* **153**, 13-29 (1981).
- Singer, A., Hathcock, K. S. & Hodes, R. J. *J. exp. Med.* **153**, 1286-1301 (1981).
- Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49**, 273-280 (1987).
- Kappler, J. W., Staerz, U., White, J. & Marrack, P. C. *Nature* **322**, 35-40 (1988).
- MacDonald, H. R. *et al.* *Nature* **332**, 40-45 (1988).
- Kisielow, P. *et al.* *Nature* **333**, 742-746 (1988).
- Fowlkes, B. J., Schwartz, R. H. & Pardoll, D. M. *Nature* **334**, 620-623 (1988).
- Von Boehmer, H., Teh, M. S., Bunnick, J. R. & Haas, W. *Recognition and Regulation in Cell-mediated Immunity* (eds Watson, J. D. & Marbrook, J.) 89-106 (Marcel Dekker Inc., New York and Basel, 1985).
- Singer, A. *J. Immunol.* **140**, 2481-2483 (1988).
- Marusic-Galesic, S., Stephany, D. A., Longo, D. L. & Kruisbeek, A. M. *Nature* **333**, 180-183 (1988).
- Kappler, J. W. *et al.* *Cell* **49**, 263-271 (1987).
- Raulet, D. H. *Cell* **49**, 153-154 (1987).
- Ozato, K., Mayer, N. & Sachs, D. H. *Transplantation* **34**, 113-118 (1982).
- Ozato, K., Mayer, N. & Sachs, D. H. *J. Immunol.* **124**, 533-540 (1980).
- Marusic-Galesic, S., Longo, D. L. & Kruisbeek, A. M. *J. exp. Med.* (in the press).
- Springer, T., Galfre, G., Secher, D. S. & Milstein, C. *Eur. J. Immunol.* **8**, 539-547 (1978).
- Crispe, I. N. & Bevan, M. J. *J. Immunol.* **138**, 2013-2018 (1987).
- Fowlkes, B. J. & Mathieson, B. J. *Surv. Immunol. Res.* **4**, 96-109 (1985).
- Marrack, P., Kuschner, P., Born, W., McDuffie, M. & Kappler, J. *J. Immunol.* **140**, 2508-2514 (1988).
- Teh, H. S. *et al.* *Nature* **335**, 229-233 (1988).
- Kisielow, P., Teh, H. S., Bluthmann, H. & von Boehmer, H. *Nature* **335**, 730-733 (1988).
- Shi, W. C. *et al.* *Nature* **335**, 271-274 (1988).
- Pullen, A. M., Marrack, P. & Kappler, J. W. *Nature* **335**, 796-801 (1988).
- Fry, A. M. & Matis, L. A. *Nature* **335**, 830-832 (1988).
- Abe, R., Vacchio, M. S., Fox, B. & Hodes, R. J. *Nature* **335**, 827-829 (1988).
- Marrack, P. & Kappler, J. W. *Nature* **332**, 840-842 (1988).
- Marrack, P. *et al.* *Cell* **53**, 627-634 (1988).
- Bluestone, J. A., Pardoll, D., Sharrow, S. O. & Fowlkes, B. J. *Nature* **326**, 82-84 (1987).
- Dialynas, D. P. *et al.* *Immunol. Rev.* **74**, 29-56 (1983).

ACKNOWLEDGEMENTS. We thank Drs Louis A. Matis, Ronald N. Germain and David H. Raulet for discussions and critical review of the manuscript; Fran Hausman and David A. Stephany for FCM analysis, and Ellie Irbay and Crystal Swanson for preparation of the manuscript. J.C.Z.-P. performed this work in partial fulfillment of the Ph.D. requirements at the Genetics Department, George Washington University, Washington DC.

FIG. 1 Nucleotide sequence and predicted amino acid sequence of rat brain guanylate cyclase cDNA. Nucleotides and amino acids are numbered on the left. Putative signal and transmembrane sequences are underlined, as is the polyadenylation signal. Six potential N-glycosylation sites in the putative extracellular domain are indicated by dashed lines under the amino-acid sequences. The six cysteine residues in the putative extracellular domain are boxed, as are the initiation codon, two upstream ATG sequences and four upstream stop codons.

METHODS. An oligo(dT)-primed, size-selected (>2 kilobases) library was constructed in phage λ gt11, using mRNA from adult rat brain. Partial-length guanylate cyclase cDNA clones were isolated from human cDNA libraries (D.G.L. *et al.*, submitted) using a sea urchin guanylate cyclase cDNA probe⁶. These partial-length cDNA inserts were then used as probes to screen the amplified rat brain λ gt11 library, grown in *Escherichia coli* Y1088 cells. Plaques were transferred to nylon filters and hybridized to random-primed, radiolabelled probes in 2 $\times$ SSC, 0.1% SDS, 5 $\times$ Denhardt's solution, 100 μ g ml⁻¹ salmon sperm DNA, at 65 $^{\circ}$ C; filters were washed with 2 $\times$ SSC, 0.1% SDS at 65 $^{\circ}$ C. The longest clone obtained was chosen for sequence analysis. The two cDNA fragments obtained by *Eco*RI digestion of the λ gt11 clone were subcloned into the *Eco*RI site of Bluescript KS⁺ (Stratagene) for restriction mapping, then the appropriate restriction fragments were subcloned into M13mp18 and M13mp19 for dideoxy sequencing using Sequenase (United States Biochemical). In addition, deletions were made from the 3' end of the 5' 1.2-kilobase *Eco*RI fragment in Bluescript, as a portion of this fragment lacked convenient restriction sites for subcloning into M13. The Erase-a-Base kit (Promega) was used for the exonuclease III/S1 nuclease treatments. Single-stranded templates for sequencing were rescued from the deletion clones in Bluescript using M13K07 as helper phage.

1 AAGCGCTCTCGGCTCTCGGACGCTCCCAATTACGCTCTGCTCGACGGCGAACCCTCGCAGCCTCCGACGGCAGCGTCCCTCGGGTTGCGGCTTCAACCCACCCGAGTTCTCCCTCGTACGACTCGGGGCGCCCTGGACGTTTCGACCCCTCGC

160 CGCTGAGCCCGAGGATGCGGACGAGCAATGTCACAGCGCTGCCGGTCTGCTGCACTCGCTGAGGCC ATG CCG GGC TCC CGA CGC GTC CGT CCG CGC CTA AGG GCG CTG CTG CTG CCG CCG CTT CTG CTA
-28 Met Pro Gly Ser Arg Arg Val Arg Pro Arg Leu Arg Ala Leu Leu Leu Leu Pro Pro Leu Leu Leu

294 CTC CGG GGC GGC CAC GCG AGC GAC CTG ACC GTG GCT GTG GTG CTG CCG CTG ACC AAC ACC TCG TAC CCG TGG TCC TGG GCG CGT GTA GGG CCG GCC GTG GAA CTG GCT CTC GCG CGG GTG
-6 Leu Arg Gly Gly His Ala Ser Asp Leu Thr Val Ala Val Val Leu Pro Leu Thr Asn Thr Ser Tyr Pro Trp Ser Trp Ala Arg Val Gly Pro Ala Val Gly Leu Ala Leu Ala Arg Val

414 AAG GGT CCG CCG GAC TTG CTG CCG GGT TGG ACG GTG CCG ATG GTG CTG GGC AGC AGT GAG AAC CCG GCG GGC GTC TGC TCG GAC ACC GCC GCA CCG CTG GCC GCG GTG GAC CTC AAG TGG
35 Lys Ala Arg Pro Asp Leu Leu Pro Gly Trp Thr Val Arg Met Val Leu Gly Ser Ser Glu Asn Ala Ala Gly Val Cys Ser Asp Thr Ala Ala Pro Leu Ala Ala Val Asp Leu Lys Trp

534 GAG CAC AGC CCC GCG GTG TTC CTG GGC CCC GGC TGC GTC TAC TCC GCT GCC CCG GTG GGG CCG TTC ACC GCG CAC TGG CCG GTG CCG CTG CTG ACC GCC GGC GCC CCG GCT CTG GGC ATC
75 Glu His Ser Pro Ala Val Phe Leu Gly Pro Gly Cys Val Tyr Ser Ala Ala Pro Val Gly Arg Phe Thr Ala His Trp Arg Val Pro Leu Leu Thr Ala Gly Ala Pro Ala Leu Gly Ile

654 GGG GTC AAG GAT GAG TAT GCG CTA ACC ACC CCG ACA GGA CCC ACC CAT GTC AAG CTG GGC GAT TTC GTG ACG GCG CTG CAT CGA CCG CTG GGC TGG GAG CAC CAG GCG CTG GTG CTC TAT
115 Gly Val Lys Asp Glu Tyr Ala Leu Thr Thr Arg Thr Gly Pro Ser His Val Lys Leu Gly Asp Phe Val Thr Ala Leu His Arg Arg Leu Gly Trp Glu His Gln Ala Leu Val Leu Tyr

774 GCA GAT CCG CTG GGC GAC GAC CCG CCT TGC TTC TTA ATA GTG GAG GGG CTG TAC ATG CCG GTG CGT GAA CGC CTC AAC ATC ACA GTG AAT CAC CAG GAG TTC GTC GAG GGC GAC CCG GAC
155 Ala Asp Arg Leu Gly Asp Asp Arg Pro Cys Phe Phe Ile Val Glu Gly Leu Tyr Met Arg Val Arg Glu Arg Leu Asn Ile Thr Val Asn His Gln Glu Phe Val Glu Gly Asp Pro Asp

894 CAC TAC CCC AAG CTA CTG CCG GCC GTG CCG CGA AAG GGC AGA GTT ATC TAC ATC TGC AGT TCT CCG GAT GCC TTC AGG AAT CTG ATG CTT CTG GCC CTG AAC GCT GGC CTG ACT GGG GAG
195 His Tyr Pro Lys Leu Leu Arg Ala Val Arg Arg Lys Gly Arg Val Ile Tyr Ile Cys Ser Ser Pro Asp Ala Phe Arg Asn Leu Met Leu Leu Ala Leu Asn Ala Gly Leu Thr Gly Glu

1014 GAC TAT GTT TTC TTA CAC CTG GAT GTG TTT GGG CAA AGC CTT AAG AGT GCT CAG GGC CTT GTT CCC CAG AAA CCC TGG GAA AGA GGA GAT GGG CAG GAC AGG AGT GCC CCG CAG GCC TTT
235 Asp Tyr Val Phe Phe His Leu Asp Val Phe Gly Gln Ser Leu Lys Ser Ala Gln Gly Leu Val Pro Gln Lys Pro Trp Glu Arg Gly Asp Gly Gln Asp Arg Ser Ala Arg Gln Ala Phe

1134 CAG GCT GCC AAA ATT ATT ACT TAC AAA GAG CCT GAT AAT CCT GAG TAC TTG GAA TTC CTG AAG CAG CTG AAA CTC TTG GCT GAC AAG AAG TTC AAC TTC ACC GTG GAG GAT GGC CTG AAG
275 Gln Ala Ala Lys Ile Ile Thr Tyr Lys Glu Pro Asp Asn Pro Glu Tyr Leu Glu Phe Leu Lys Gln Leu Lys Leu Leu Ala Asp Lys Lys Phe Asn Phe Thr Val Glu Asp Gly Leu Lys

1254 AAT ATC ATC CCA GCC TCC TTC CAC GAC GGG CTC CTG CTC TAT GTC CAG GCA GTG ACA GAG ACT CTG GCA CAG GGG GGA ACT GTC ACA GAT GGA GAG AAC ATC ACT CAG CCG ATG TGG AAC
315 Asn Ile Ile Pro Ala Ser Phe His Asp Gly Leu Leu Leu Tyr Val Gln Ala Val Thr Glu Thr Leu Ala Gln Gly Gly Thr Val Thr Asp Gly Glu Asn Ile Thr Gln Arg Met Trp Asn

1374 CGA AGC TTC CAA GGT GTG ACA GGA TAC CTG AAA ATT GAT AGA AAC GGA GAT CCG GAC ACC GAT TTC TCT CTC TGG GAT ATG GAT CCA GAG ACG GGT GCC TTC AGG GTT GTC CTG AAC TAT
355 Arg Ser Phe Gln Gly Val Thr Gly Tyr Leu Lys Ile Asp Arg Asn Gly Asp Arg Asp Thr Asp Phe Ser Leu Trp Asp Met Asp Pro Trp Gly Ala Phe Arg Val Val Leu Asn Tyr

1494 AAT GGT ACT TCC CAG GAG CTA ATG GCT GTG TCA GAA CAC AAA TTA TAC TGG CCT CTG GGA TAT CCA CCT CCT GAC GTC CCT AAA TGT GGC TTT GAC AAT GAG GAC CCA GCC TGC AAC CAA
395 Asn Gly Thr Ser Gln Glu Leu Met Ala Val Ser Glu His Lys Leu Tyr Trp Pro Leu Gly Tyr Pro Pro Pro Asp Val Pro Lys Cys Gly Phe Asp Asn Glu Asp Pro Ala Cys Asn Gln

1614 GAC CAC TTT TCC ACA CTG GAG GTT CTG GCT TTG GTG GGC AGC CTC TCT GTT AGC TTT CTG ATT GTG TCT TTC TTA TAC AGG AAG ATG CAG CTG GAA AAG GAG CTG GTC TCA GAG
435 Asp His Phe Ser Thr Leu Glu Val Leu Ala Leu Val Gly Ser Leu Ser Leu Ile Ser Phe Leu Ile Val Ser Phe Phe Ile Tyr Arg Lys Met Gln Leu Glu Lys Glu Leu Val Ser Glu

1734 TTG TGG CCG GTG CCG TGG GAG GAC TTG CAG CCC AGC AGC CTG GAG AGG CAT CTT CCG AGC GCT GGC AGC CCG CTG ACC CTG AGT GGG CGA GGC TCC AAT TAT GGC TCC CTG CTA ACC ACC
475 Leu Trp Arg Val Arg Trp Glu Asp Leu Gln Pro Ser Ser Leu Glu Arg His Leu Arg Ser Ala Gly Ser Arg Leu Thr Leu Ser Gly Arg Gly Ser Asn Tyr Gly Ser Leu Leu Thr Thr

1854 GAG GGC CAG TTC CAA GTC TTT GCC AAG ACA GCA TAC TAT AAG GGC AAC CTT GTG GCT GTG AAA CGT GTG AAC CCG AAA CCG ATT GAG TTG ACA CGA AAA GTC CTG TTT GAA CTT AAA CAT
515 Glu Gly Gln Phe Gln Val Phe Ala Lys Thr Ala Tyr Tyr Lys Gly Asn Leu Val Ala Val Lys Arg Val Asn Arg Lys Arg Ile Glu Leu Thr Arg Lys Val Leu Phe Glu Leu Lys His

1974 ATG CCG GAT GTG CAG AAT GAG CAC TTG ACA AGA TTT GTG GGT GCT TGT ACC GAC CCC CCC AAC ATC TGT ATC CTC ACA GAG TAC TGT CCC CGT GGA AGC CTA CAG GAC ATT CTA GAG AAT
555 Met Arg Asp Val Gln Asn Glu His Leu Thr Arg Phe Val Gly Ala Cys Thr Asp Pro Pro Asn Ile Cys Ile Leu Thr Glu Tyr Cys Pro Arg Gly Ser Leu Gln Asp Ile Leu Glu Asn

2094 GAG AGT ATC ACC CTG GAC TGG ATG TTT CCG TAC TCG CTC ACC AAT GAC ATT GTC AAG GGA ATG CTC TTT CTA CAC AAT GGG GCC ATT TGT TCC CAT GGG AAC CTC AAG TCA TCC AAC TGT
595 Glu Ser Ile Thr Leu Asp Trp Met Phe Arg Tyr Ser Leu Thr Phe Ser Leu Thr Phe His Lys Asn Gly Ala Ile Cys Ser His Gly Asn Leu Lys Ser Ser Asn Cys

2214 GTG GTA GAC GGG CCG TTC GTG TTA AAG ATC ACA GAC TAC GGT CTT GAG AGC TTC AGA GAC CCG GAG CCA GAG CAA GGG CAC ACC CTC TTT GCC AAA AAA TTG TGG ACG GCA CCT GAG CTC
635 Val Val Asp Gly Arg Phe Val Leu Lys Ile Thr Asp Tyr Gly Leu Glu Ser Phe Arg Asp Pro Glu Pro Glu Gln Gly His Thr Leu Phe Ala Lys Lys Leu Trp Thr Ala Pro Glu Leu

2334 CTG CGA ATG GCT TCG CCA CCT GCC CGT GGC TCC CAA GCT GGG GAT GTG TAC AGC TTT GGT ATC ATC CTG CAG GAG ATT GCC CTA AGA AGT GGG GTC TTC TAT GTG GAA GGT TTG GAC CTC
675 Leu Arg Met Ala Ser Pro Pro Ala Arg Gly Ser Gln Ala Gly Asp Val Tyr Ser Phe Gly Ile Ile Leu Gln Glu Ile Ala Leu Arg Ser Gly Val Phe Tyr Val Glu Gly Leu Asp Leu

2454 AGC CCA AAA GAG ATC ATT GAG CGT GTG ACT CCG GGT GAG CAG CCC CCA TTC CGA CCC TCC ATG GAT CTG CAG AGC CAC CTG GAG GAA CTG GGG CAG CTG ATG CAG CCG TGC TGG GCA GAG
715 Ser Pro Lys Glu Ile Ile Glu Arg Val Thr Arg Gly Glu Gln Pro Pro Phe Arg Pro Ser Met Asp Leu Gln Ser His Leu Glu Glu Leu Gly Gln Leu Met Gln Arg Cys Trp Ala Glu

2574 GAC CCA CAG GAG CCG CCA CCC TTT CAG CAG ATC CCG CTG GCG CTG CCG AAG TTC AAC AAG GAG AAC AGC AGC AAC ATC CTG GAC AAC CTG CTG TCA CCG ATG GAG CAG TAT GCT AAC AAC
755 Asp Pro Gln Glu Arg Pro Pro Phe Gln Gln Ile Arg Leu Ala Leu Arg Lys Phe Asn Lys Glu Asn Ser Ser Asn Ile Leu Asp Asn Leu Leu Ser Arg Met Glu Gln Tyr Ala Asn Asn

2694 CTG GAG GAA CTG GTA GAG GAG AGA ACA CAA GCT TAT CTG GAG GAG AAG CCG AAA GCT GAG GCC TTG CTT TAC CAG ATT CTG CCT CAC TCC GTG GCT GAG CAG CTG AAG AGA GGC GAG ACA
795 Leu Glu Glu Leu Val Glu Leu Arg Thr Gln Ala Tyr Leu Glu Glu Asn Lys Arg Lys Ala Glu Ala Leu Phe Leu Thr Gln Ile Leu Pro His Ser Val Ala Glu Gln Leu Lys Arg Gly Glu Thr

2814 GTC CAG GCT GAG GCC TTT GAT AGT GTT ACC ATC TAC TTC AGT GAT ATT GTG GGC TTT ACA GCT CTT TCA GCA GAA AGC ACA CCC ATG CAG GTG GTG ACT CTG CTC AAT GAT CTG TAC ACC
835 Val Gln Ala Glu Ala Phe Asp Ser Val Thr Ile Tyr Phe Ser Asp Ile Val Gly Phe Thr Ala Leu Ser Ala Glu Ser Thr Pro Met Gln Val Val Thr Leu Leu Asn Asp Leu Tyr Thr

2934 TGT TTT GAT GCT GTC ATA GAC AAC TTT GAT GTG TAC AAG GTG GAC ACC ATT GGT GAT GCT TAC ATG GTG GTG TCA GGG CTC CCA GTG CCG AAT GGA CAA CTC CAC GCC CGA GAG GTG GCC
875 Cys Phe Asp Ala Val Ile Asp Asn Phe Asp Val Tyr Lys Val Glu Thr Ile Gly Asp Ala Tyr Met Val Val Ser Gly Leu Pro Val Arg Asn Gly Gln Leu His Ala Arg Glu Val Ala

3054 CGA ATG GCA CTT GCA CTA CTG GAT GCT GTG CCG TCC TTC CCG CAT AGG CCC CAG GAA CAG CTG CCG TTG CCG ATT GGC ATC CAC ACA GGT CCT GTG TGT GCT GGT GTG GTA GGG
915 Arg Met Ala Leu Ala Leu Leu Asp Ala Val Arg Ser Phe Arg Ile Arg His Arg Pro Gln Glu Gln Leu Arg Leu Arg Ile Gly Ile His Thr Gly Pro Val Cys Ala Gly Val Val Gly

3174 CTA AAG ATG CCC CGA TAC TGC CTC TTT GGA GAC ACA GTC AAC ACA GCT TCA AGA ATG GAG TCT AAT GGA GAA GCC CTC AAG ATC CAC TTG TCT TCA GAG ACC AAG GCT GTG CTG GAA GAG
955 Leu Lys Met Pro Arg Tyr Cys Leu Phe Gly Asp Thr Val Asn Thr Ala Ser Arg Met Glu Ser Asn Gly Glu Ala Leu Lys Ile His Leu Ser Ser Glu Thr Lys Ala Val Leu Glu Glu

3294 TTC GAT GGT TTC GAG CTG GAG CTC CGA GGG GAT GTG GAA ATG AAG GGC AAA GGC AAG GTT CCG ACC TAT TGG CTC CTG GGG GAG CCG GGA TGT AGC ACT CGA GGC TGA CTTACTGCCCTG
995 Phe Asp Gly Glu Glu Leu Glu Arg Val Glu Met Lys Gly Lys Gly Lys Val Arg Thr Tyr Trp Leu Leu Gly Glu Arg Gly Cys Ser Thr Arg Gly *

3414 CTGTTCTCTGTCACCCCTCCTCCCTGTGCCAGAGGTGACAGAGGTGCCAGCTTCCACCTCTCCACAGCAGCCGACCTGTGGAAGGATTAGGACCTGACCAGCAGTACCAGATGTGACCTCTGAGAGAGGATGGAGATGGTGGGACTGCA

3573 GGGGACCACTAAGTTTGTAGGACTGACTGAAACACAGTCCCTCCCATGGCACCCCTGTGGCCACACATGCCAGTCCACCCCTTACTCTGCTGCCTAGATTGGGACAGGATTCTTCTCTGCCCTCAACTAGCTCCACTGTGACTTATAGGGAGGGAA

3732 TTGCCACCTGAAGGAACAGAAAGAGGTTAGAGTTTGACGAGGCGAGGCCTCTGTGTACAAATACTCCCTCACTTCCAGCCCAACCTGCCCCACAGACTTGGACACAGCTCACTGAGGAGAGAGAGAGCTGCCGGTACCTTGTCTCTCGTG

3891 TGAACCAACCAATTAAGTCTTTATTCTGTGTGAAAAAATAA

FIG. 2 Alignment of deduced amino-acid sequences of the membrane form of guanylate cyclase from rat brain with a, the bovine ANP-C receptor¹⁶; b, the murine PDGF receptor¹⁷; c, bovine soluble guanylate cyclase¹⁹ and yeast adenylate cyclase²⁰; d, the membrane form of guanylate cyclase from the sea urchin *Arbacia punctulata*⁶. The entire sequence of the mature rat guanylate cyclase is shown in a-c. Boxed residues represent the putative transmembrane domains (a), the 33 residues most highly conserved among protein kinases¹⁸ (b), and sequences extending beyond the kinase-like domain (d). Regions of identity are indicated by asterisks.

human cDNA libraries (D.G.L. *et al.*, manuscript in preparation). These clones were used in turn, to probe a size-selected rat brain cDNA library to isolate a cloned cDNA containing the entire coding region of guanylate cyclase.

The nucleotide and deduced amino-acid sequences of the guanylate cyclase cDNA are shown in Fig. 1. An open reading frame encoding 1,057 amino acids is preceded by an in-frame stop codon and is flanked by 227 base pairs (bp) and 535 bp of 5' and 3' untranslated sequences respectively. Hydrophobic analysis predicts an amino-terminal signal sequence and a single transmembrane domain that divides the protein into a 441-amino-acid extracellular domain and a 567-amino-acid intracellular domain. Using an algorithm for prediction of the signal cleavage site⁹, residue 29 is expected to represent the amino terminus of the mature protein, which would have a relative molecular mass (M_r) of 115,852 and an isoelectric point of 6.33. Taking into account post-translational processing known to occur in guanylate cyclase^{10,11}, the predicted M_r is consistent with estimates of 120,000–180,000 (120K–180K) for membrane forms of guanylate cyclase reported in the literature^{10–13}. Six cysteine residues and six potential N-glycosylation sites are present in the extracellular domain.

Further analysis of the deduced amino-acid sequence of the guanylate cyclase cDNA indicates that it may be divided into three potential functional domains (Fig. 2a–c). The extracellular domain of the rat cell membrane guanylate cyclase is 33% identical with the bovine ANP-clearance (ANP-C) receptor, an ANP-binding protein that is apparently not coupled to activation of guanylate cyclase or to the biological effects of ANP (refs 14–16) (Fig. 2a). The ANP-C receptor consists of an extracellular ANP-binding domain, a transmembrane domain and a short cytoplasmic tail¹⁶; it has been suggested that the function of this protein is to clear excess ANP from the circulation¹⁴. The five cysteine residues of the ANP-C receptor are conserved in guanylate cyclase.

As noted earlier for guanylate cyclase from sea urchin⁶, the intracellular domain is related to the catalytic domain of protein kinases, although protein kinase activity has not yet been detected. A 256-amino acid portion of the intracellular domain is 31% identical to the protein tyrosine kinase domain of the platelet-derived growth factor (PDGF) receptor¹⁷ (Fig. 2b). Guanylate cyclase conforms to the protein kinase consensus sequence¹⁸ in 30 out of 33 residues. Exceptions are Ser⁵⁰² (Leu, Ile or Val in all known protein kinases), Asn⁶²⁸ (Asp in all known protein kinases), and Ala⁶⁶⁵ (Ser, Thr or Pro at this position in 63 out of 65 protein kinases). In addition, the Gly-X-Gly-X-X-Gly consensus sequence of protein kinases is Gly-X-Gly-X-X-X-Gly in the rat guanylate cyclase.

The most extensive similarity between the deduced amino-acid sequence of the membrane form of guanylate cyclase and other proteins is found in the carboxyl portion of the intracellular domain. A 253-amino acid sequence in this region is 42% identical to the carboxyl terminus of one of the subunits of a bovine soluble form of guanylate cyclase¹⁹ (Fig. 2c). It is not yet known whether the sequence determined for the soluble enzyme is from a regulatory or catalytic subunit¹⁹. But only a 77-residue portion of this 253-amino acid region is conserved in guanylate cyclase from the sea urchin *Arbacia punctulata* (see below), which implies that it may not be the guanylate cyclase

catalytic domain. This region of the carboxyl tail has limited (23%) sequence identity with a 193-amino-acid portion of the yeast adenylate cyclase²¹ (Fig. 2c); the significance, if any, of this limited similarity is not known.

The deduced amino-acid sequence of rat brain guanylate cyclase is 45% identical to that of the membrane form of guanylate cyclase from the sea urchin *Arbacia punctulata*⁶ within a 399-amino acid region of the intracellular domain beginning immediately after the transmembrane domain (Fig. 2d). This region consists primarily of the protein kinase-like domain but also includes 77 amino acids carboxyl to this domain (Fig. 2d). Presumably this highly conserved region includes the guanylate cyclase catalytic domain, whereas the divergent regions may represent regulatory domains.

To confirm that the isolated cDNA encoded a protein with guanylate cyclase activity, we inserted the cDNA in a mammalian expression vector (pSVL) under the control of the SV40 late promoter. Because upstream AUG codons have been reported to reduce the efficiency of translation of messenger

TABLE 1 Particulate guanylate cyclase activity and ¹²⁵I-labelled ANP-binding activity of transfected COS-7 cells

Transfection	Guanylate cyclase activity (pmol cGMP min ⁻¹ per mg protein)		Specific ¹²⁵ I-ANP binding (c.p.m.)
	–ANP	+ANP	
pSVL	0.38	1.1	3,082
pSVL-GC	16.3	55.8	28,039

COS-7 cells, maintained in Dulbecco's modified Eagle's medium (DME) supplemented with 20 mM HEPES at pH 7.4 with penicillin/streptomycin and 10% fetal bovine serum, were transfected with plasmid DNA using DEAE-dextran²⁸. Cells were seeded at 5×10^5 cells per 100-mm plate (guanylate cyclase assay) or 5×10^5 cells per 35-mm plate (¹²⁵I-labelled ANP binding assay). Cells were fed the day after transfection and assayed for guanylate cyclase activity or ¹²⁵I-ANP binding two days after transfection. The guanylate cyclase construct in pSVL (Pharmacia) was prepared as follows. The 5' 1.2-kb *EcoRI* fragment of the guanylate cyclase cDNA was cloned into the *EcoRI* site of Bluescript KS⁺, oriented so that the 5' end of the insert was closest to the *HindIII* site in the polylinker. The 194 bp at the 5' end of the insert were removed by digestion with *HindIII* and *NcoI*, and the 4-kb fragment containing the remainder of the cDNA and the vector was purified by agarose gel electrophoresis and adsorption to glass beads. The DNA was then treated with *S1* nuclease and religated. The extent of the resulting deletion was determined by double-stranded sequencing. The 3' 2.7-kb *EcoRI* fragment of the guanylate cyclase cDNA was then cloned into the *EcoRI* site of the modified plasmid, and its orientation confirmed by restriction analysis. The *SalI/SmaI* fragment containing the guanylate cyclase cDNA with the upstream ATGs removed was then cloned into the *XhoI/SmaI* site of pSVL to create pSVL-GC. For guanylate cyclase assays, 100-mm cultures on ice were washed twice with 5 ml cold HEPES-buffered saline, scraped into 3 ml cold 20 mM HEPES, pH 7.4, 50 mM NaCl, 5 mM EDTA, 1 mM dithiothreitol, and homogenized by passage ten times through a 22-gauge needle. Following centrifugation for 15 min at 5,000g and washing with the same buffer, membrane proteins were solubilized by incubation on ice for 30 min in a solution containing 20 mM HEPES, pH 7.4, 100 mM NaCl, 10% glycerol, 1% Triton X-100, 1 mM dithiothreitol. After centrifugation for 15 min at 5,000g, supernatant fluids were adjusted to equal protein concentrations and incubated for 10 min in the absence or presence of 1 μ M ANP at 0 °C. Incubation was continued for 30 min at 30 °C after diluting 100 μ l each sample into a solution containing a final concentration of 20 mM HEPES, pH 7.4, and 1 mM each GTP, MnCl₂ and methyl isobutylxanthine, in a total reaction volume of 250 μ l. Reactions were stopped by the addition of 750 μ l 0.5 M perchloric acid and samples were analysed for cGMP by radioimmunoassay after purification of nucleotides by Dowex chromatography²⁶. ¹²⁵I-labelled ANP binding was measured essentially as described¹⁵. Cells in 35-mm plates were washed twice with 2 ml DME/0.1% bovine serum albumin (BSA), then incubated in 0.5 ml the same buffer containing 0.5 nM ¹²⁵I-labelled rat ANP (2,200 Ci mmol⁻¹, NEN) in the absence or presence of 0.5 μ M unlabelled rat ANP (Sigma) for 20 min at 37 °C. After five 1-ml washes with HBSS/0.1% BSA, cells were solubilized in 1 ml 0.5 M NaOH and the radioactivity determined. All determinations were in triplicate (nonspecific binding was ~5,000 c.p.m. in each case).

RNA^{21,22}, 194 bp of 5' untranslated sequences were removed from the cDNA before its insertion into pSVL; this resulted in the deletion of the two ATG sequences 5' to the initiation site (see Fig. 1 and Table 1 legend). This modification of the cDNA was found to enhance expression of guanylate cyclase about tenfold in transfected mammalian cells (data not shown). COS-7 cells were transfected with the guanylate cyclase expression plasmid (pSVL-GC) or with pSVL, and a particulate fraction was isolated from the cells and assayed for enzyme activity. The activity of membrane-associated guanylate cyclase was increased by about 40-fold in cells transfected with pSVL-GC (Table 1), which is consistent with the clone encoding guanylate cyclase. ANP stimulated guanylate cyclase activity about threefold in detergent extracts of both control cells and cells expressing the cloned guanylate cyclase (Table 1).

The cloned guanylate cyclase also functions as an ANP-receptor (Table 1). Cells transfected with pSVL-GC specifically bound about nine times more ¹²⁵I-labelled ANP than cells transfected with pSVL alone. ¹²⁵I-labelled ANP bound with high affinity and the expected specificity (half-maximal inhibition of binding of 0.4 nM ¹²⁵I-labelled ANP occurred at about 3 nM unlabelled ANP or 100 nM unlabelled atriopeptin I; data not shown). To verify that guanylate cyclase is acting as an ANP-receptor, we used ¹²⁵I-labelled ANP in cross-linking experiments (Fig. 3). Transfected cells were incubated with ¹²⁵I-labelled ANP with or without excess unlabelled ANP. After cross-linking and SDS-polyacrylamide gel electrophoresis under reducing conditions, a single band migrating with *M_r* 130K was evident from

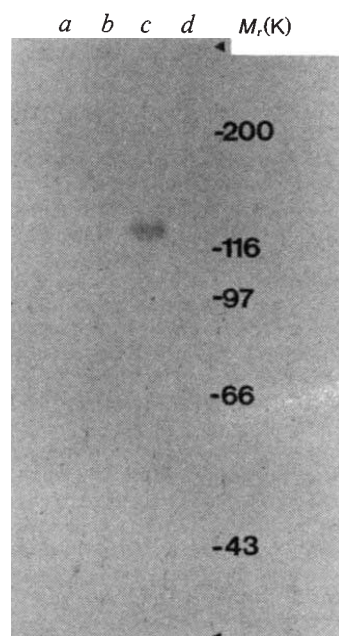


FIG. 3 Cross-linking of ¹²⁵I-labelled ANP to transfected COS-7 cells. Transfections are described in the legend to Table 1. Binding and cross-linking have been described previously¹⁵. Cells were washed 3 times with 2 ml Hanks' balanced salt solution (HBSS), then incubated for 1 h at room temperature in 1 ml HBSS containing 0.5 nM ¹²⁵I-labelled ANP (2,200 Ci mmol⁻¹) in the absence (a, c) or presence (b, d) of 0.5 μM unlabelled ANP. 1 ml HBSS containing 2 mM disuccinimidyl suberate was added to a final concentration 1 mM and the incubation continued for 30 min. Cells were then washed with 2 ml HBSS and solubilized in 400 μl hot (95 °C) Laemmli sample buffer²⁹, of which 150 μl was heated at 95 °C in the presence of 7.5 μl 2-mercaptoethanol and analysed by electrophoresis on a 7.5% SDS-polyacrylamide gel²⁹, followed by drying and autoradiography. Positions of molecular weight markers are shown on the right; the top and bottom of the separating gel are indicated by arrowheads. a, b, Cells transfected with pSVL; c, d, cells transfected with pSVL-GC.

the cells transfected with pSVL-GC. This band was not seen in the presence of an excess of non-radioactive ANP or in cells transfected with pSVL. The mobility of this radioactive band coincides with that previously reported for the high-molecular weight ANP receptor^{12,15,23,24}. Thus the guanylate cyclase cDNA encodes a protein having both guanylate cyclase and ANP-binding activities.

In conclusion, we have isolated and sequenced a cDNA clone for the membrane form of guanylate cyclase from rat brain. From the cDNA sequence we predict a protein divided by a single transmembrane domain into an amino-terminal extracellular ANP-binding domain, and a carboxyl-terminal intracellular catalytic domain (Fig. 4). The intracellular domain is homologous both with protein kinases and to a subunit of the soluble form of guanylate cyclase, whereas the extracellular domain is homologous with a low-molecular weight ANP receptor. Expression of the cDNA produces a protein with both ANP-binding and guanylate cyclase activity.

Previous cross-linking experiments indicated that the membrane form of guanylate cyclase in *Arbacia punctulata* could function as a receptor for egg-associated peptides which activate sperm respiration and motility²⁵; one of the earliest effects of these egg peptides on spermatozoa is the stimulation of cyclic GMP synthesis^{26,27}. But functional expression of the cloned sea urchin enzyme has not yet been obtained in mammalian cells or in *Xenopus* oocytes⁶, apparently because of a lack of proper post-translational modification. It has not been possible, therefore, to test directly whether the enzyme serves as a receptor. Several groups have also reported the co-purification of high-molecular weight ANP receptors with active guanylate cyclase^{12,13,24}, but it is not clear whether the two activities reside within the same molecule¹². Here we have demonstrated directly that the membrane form of guanylate cyclase functions as a receptor. The ANP receptor/guanylate cyclase we describe may represent, by analogy with the growth factor receptor/protein

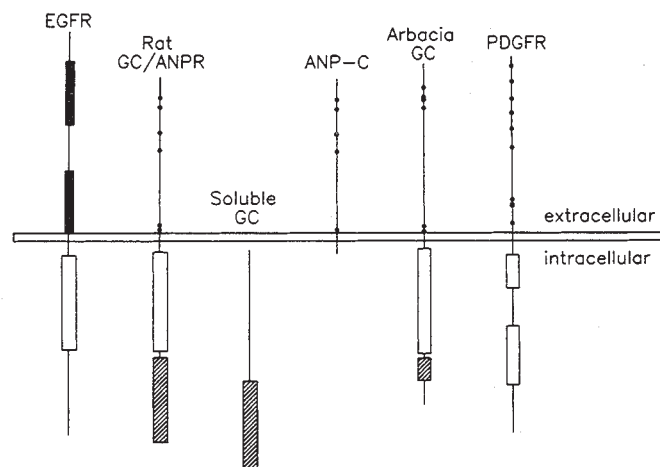


FIG. 4 Topology of members of the guanylate cyclase/ANP receptor family and the growth factor receptor/protein tyrosine kinase family. Drawn to scale are representations of the epidermal growth factor receptor (EGFR)³⁰, the membrane forms of guanylate cyclase from rat (rat GC/ANPR) and from the sea urchin *Arbacia punctulata* (Arbacia GC)⁶, a subunit of the soluble form of guanylate cyclase (soluble GC)¹⁹, the ANP-C receptor (ANP-C)¹⁸, and the platelet-derived growth factor receptor (PDGFR)¹⁷. Drawing modelled after ref. 17. Individual cysteine residues are indicated by black circles in the extracellular domains of each protein, except in the case of the EGF receptor, in which the cysteine-rich domains are indicated by black boxes. Protein kinase domains (EGFR and PDGFR) and protein kinase-like domains (rat GC/ANPR, Arbacia GC) are indicated by open boxes. Areas of sequence similarity between the soluble and membrane forms of guanylate cyclase are indicated by hatched boxes.

tyrosine kinases, a member of a diverse family of cell-surface receptors. This family is unique with regard to the regulation of cyclic nucleotides, in that cGMP is formed directly by the receptor protein in response to the binding of an agonist. □

Received 14 December 1988; accepted 11 January 1989.

1. Flynn, T. G. & Davies, P. L. *Biochem. J.* **232**, 313–321 (1985).
2. Hamet, P. et al. *Biochem. biophys. Res. Commun.* **123**, 515–527 (1984).
3. Waldman, S. A., Rapoport, R. M. & Murad, F. *J. biol. Chem.* **259**, 14332–14334 (1984).
4. Gerzer, R., Bohme, E., Hoffman, F. & Schuit, G. *FEBS Lett.* **132**, 71–74 (1981).
5. Kamisaki, Y. et al. *J. biol. Chem.* **261**, 7236–7241 (1986).
6. Singh, S. et al. *Nature* **334**, 708–712 (1988).
7. Tremblay, J. et al. *FEBS Lett.* **181**, 17–22 (1985).
8. Winkler, R. J. et al. *Proc. natn. Acad. Sci. USA* **81**, 7661–7664 (1984).
9. Von Heijne, G. *Eur. J. Biochem.* **133**, 17–21 (1983).
10. Radany, E. W., Gerzer, R. & Garbers, D. L. *J. biol. Chem.* **258**, 8346–8351 (1983).
11. Ramarao, C. S. & Garbers, D. L. *J. biol. Chem.* **263**, 1524–1529 (1988).
12. Kuno, T. et al. *J. biol. Chem.* **261**, 5817–5823 (1986).

13. Paul, A. K., Marala, R. B., Jaiswal, R. K. & Sharma, R. K. *Science* **235**, 1224–1226 (1987).
14. Maack, T. et al. *Science* **238**, 675–678 (1987).
15. Leitman, D. C. et al. *J. biol. Chem.* **263**, 3720–3728 (1988).
16. Fuller, F. et al. *J. biol. Chem.* **263**, 9395–9401 (1988).
17. Yarden, Y. et al. *Nature* **323**, 226–232 (1986).
18. Hanks, S. K., Quinn, A. M. & Hunter, T. *Science* **241**, 42–53 (1988).
19. Koesling, D. et al. *FEBS Lett.* **239**, 29–34 (1988).
20. Kataoka, T., Broek, D. & Wigler, M. *Cell* **43**, 493–505 (1985).
21. Kozak, M. *Nucleic Acids Res.* **12**, 3873–3893 (1984).
22. Mueller, P. P. & Hinnebusch, A. G. *Cell* **45**, 201–207 (1986).
23. Yip, C. C., Laing, L. D. & Flynn, T. G. *J. biol. Chem.* **260**, 8229–8232 (1985).
24. Takayanagi, R. et al. *J. biol. Chem.* **262**, 12104–12113 (1987).
25. Shimomura, H., Dangott, L. J. & Garbers, D. L. *J. biol. Chem.* **261**, 15778–15782 (1986).
26. Hansbrough, J. R. & Garbers, D. L. *J. biol. Chem.* **256**, 1447–1452 (1981).
27. Bentley, J. K., Tubb, D. J. & Garbers, D. L. *J. biol. Chem.* **261**, 14859–14862 (1986).
28. Cullen, B. R. *Meth. Enzym.* **152**, 684–704 (1987).
29. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
30. Ulrich, A. et al. *Nature* **309**, 418–425 (1984).

ACKNOWLEDGEMENTS. We thank Rae Ann Bellet, Judy Chen, Jan Jordan, Laura Jo Rutledge, Nancy Saul and Janette Tubb for technical assistance. This work was supported in part by the NIH and by Genentech, Inc.

Molecular cloning and characterization of an insulin-regulatable glucose transporter

David E. James, Marilyn Strube & Mike Mueckler

Department of Cell Biology, Box 8101, Washington University Medical School, 660 South Euclid Ave., St Louis, Missouri 63110, USA

A MAJOR mechanism by which insulin stimulates glucose transport in muscle and fat is the translocation of glucose transporters from an intracellular membrane pool to the cell surface^{1–5}. The existence of a distinct insulin-regulatable glucose transporter was suggested by the poor cross-reactivity between antibodies specific for either the HepG2 or rat brain glucose transporters and the rat adipocyte glucose transporter^{6,7}. More direct evidence was provided by the production of a monoclonal antibody (mAb 1F8) specific for the rat adipocyte glucose transporter that immunolabels a species of relative molecular mass 43,000 (43K) present only in tissues that exhibit insulin-dependent glucose transport⁸, suggesting that this protein may be encoded by a different gene from the previously described mammalian glucose transporters^{9–13}. This antibody has been used to immunoprecipitate a 43K protein that was photoaffinity-labelled with cytochalasin B in a glucose displacable way, and to immunolabel a protein in the plasma membrane of rat adipocytes, whose concentration was increased at least fivefold after cellular insulin exposure. Here we describe the cloning and sequencing of cDNAs isolated from both rat adipocyte and heart libraries that encode a protein recognized by mAb 1F8, and which has 65% sequence identity to the human HepG2 glucose transporter⁹. This cDNA hybridizes to an mRNA present only in skeletal muscle, heart and adipose tissue. Our data indicate that this cDNA encodes a membrane protein with the characteristics of the translocatable glucose transporter expressed in insulin-responsive tissues.

Several clones were isolated from a rat adipocyte and a rat heart cDNA library by low stringency hybridization using a fragment of the HepG2 glucose transporter cDNA as the probe. A 2,086-base pair (bp) consensus sequence (the 'IRGT sequence') compiled from four independent clones obtained from both libraries is shown in Fig. 1a. This sequence includes the complete coding region for a protein that shares sequence similarity with the HepG2 glucose transporter (Fig. 1b). The initiation codon was assigned to the first ATG triplet (beginning at position 154) downstream of an in-frame termination codon (TGA) beginning at position 133. An open reading frame extends

from this ATG triplet to position 1,680 that codes for a 509-amino acid polypeptide with a predicted relative molecular mass (M_r) of 54,860. This value is in reasonable agreement with the estimated size (38K) of the deglycosylated transporter immunolabelled with mAb 1F8 (data not shown). The difference between the predicted and measured M_r values is similar to the difference between these values for the HepG2 glucose transporter, consistent with the anomalous migration of membrane proteins during SDS-PAGE⁹. Several clones were isolated from the adipocyte library that contained deletions in the 5' region of the mRNA compared to the IRGT sequence. The deleted regions were flanked by extremely GC-rich sequences. Thus, the deletions probably resulted from excision of hairpin loops formed during construction of the library.

The deduced amino-acid sequence of IRGT is 65.2% identical to that of the HepG2 glucose transporter (Fig. 1b). The sequence identity between the two proteins is particularly strong within the predicted transmembrane segments, with the notable exception of segment 3 (amino-acid residues 112–132 in the IRGT sequence, Fig. 1). Interestingly, there is little sequence homology between the IRGT (Fig. 1a), the HepG2 glucose transporter⁹ and the liver glucose transporter^{12,13} within this particular domain. There are considerable differences in the sequences of the HepG2 glucose transporter and IRGT within the predicted soluble domains of the HepG2 transporter⁹, particularly those encompassing the N and C terminus and the large, central, cytoplasmic loop. IRGT shares the putative glycosylation site (Asn 57) present in the HepG2 transporter (Asn 45) within a proposed exofacial loop after the first transmembrane segment. IRGT also contains the putative photoaffinity-labelling cytochalasin B-binding site (Trp 428) corresponding to Trp 412 in the HepG2 sequence^{14,15}. The hydropathy plot of the IRGT protein sequence, deduced using the method of Kyte and Doolittle¹⁶, is nearly superimposable on that of the HepG2 glucose transporter⁹ (data not shown), suggesting that these two proteins have very similar two-dimensional topologies with respect to the membrane. IRGT has slightly larger N-terminal, C-terminal and glycosylated loop domains compared to the HepG2 glucose transporter however.

RNA was transcribed from a 2,086-bp IRGT cDNA construction and translated in an *in vitro* rabbit reticulocyte lysate system. The primary translation product synthesized in the absence of microsomes had a slightly reduced mobility on SDS-polyacrylamide gels (apparent M_r = 39K) relative to the HepG2 translation product (apparent M_r = 37K), in reasonable agreement with the predicted difference in the relative molecular masses of the two proteins. Addition of dog pancreatic microsomes to the translation system decreased the mobility of both the IRGT (apparent M_r = 43K) and HepG2 (apparent M_r = 40K) proteins (Fig. 2) consistent with core glycosylation upon cotranslational

h

84

FIG. 1 *a*, Nucleotide sequence of IRGT cDNA and the deduced amino-acid sequence of the insulin-regulatable glucose transporter. The nucleotide sequence presented was compiled from the four clones whose sequences were identical in the overlapping regions. The two adipocyte clones encompassed residues 1–1,450 and 241–2,086. The two heart clones encompassed residues 141–1,212 and 717–1,944. Amino-acid residues are numbered above the sequence and nucleotides are numbered below. *b*, Deduced amino-acid sequences of the HepG2 glucose transporter and IRGT are aligned for comparison. The amino acids are indicated in the one-letter code. Regions of identity between the HepG2 glucose transporter and IRGT amino-acid sequences are indicated by a dash. Gaps have been introduced to optimize the alignment.

METHODS. Phage plaques (2×10^6) from both rat adipocyte and rat heart cDNA libraries in λ gt11 were screened under low stringency hybridization conditions (39 °C in 50% formamide, $5 \times$ SSPE, 0.1% SDS, $5 \times$ Dehnhardt's 250 μ g ml⁻¹ salmon sperm DNA) using a ³²P-labelled *Eco*R1 fragment (from pGT25S ref. 9) of the HepG2 glucose transporter. Filters were washed at room temperature in $1 \times$ SSC, 0.5% SDS and then at 42 °C in $0.1 \times$ SSC, 0.5% SDS. Positive clones were plaque-purified and the cDNA inserts were sized after *Eco*R1 digestion of the mini-prep DNA. Four clones (two from each library) were selected for further characterization. The inserts were subcloned into the *Eco*R1 sites of M13mp18 and mp19 and were sequenced in both strands according to a modified version²⁴ of the Sanger method²⁵. The sequence data were compiled by the computer-assisted method of Staden²⁶. The insert DNAs were sequenced completely.

insertion of the protein into microsomes. Subsequent digestion with endoglycosidase H indicated that IRGT and the HepG2 glucose transporter were core-glycosylated to the same extent during insertion into microsomes (data not shown). The apparent M_r of the core glycosylated translation product of IRGT (Fig. 2) was similar to the M_r of the protein in fat and muscle immunolabelled by mAb 1F8 (Fig. 3). The IRGT translation product, but not that of the HepG2 glucose transporter RNA, was specifically immunoprecipitated using mAb 1F8 (Fig. 2). A polyclonal antibody (R820) was prepared by immunizing rabbits with a peptide corresponding to residues 498–509 within the predicted sequence of IRGT (see Fig. 1*a*). This antiserum also specifically immunoprecipitated the IRGT translation product (Fig. 2). Polyclonal and monoclonal antibodies raised against the human erythrocyte glucose transporter immunoprecipitated the HepG2 glucose transporter translation product but not that of IRGT (data not shown).

In contrast to the tissue distribution of the HepG2 glucose transporter¹⁷ or human liver-like glucose transporter¹⁰ mRNAs, IRGT mRNA was detected at relatively high levels in brown and white adipose tissue, heart and red and white skeletal muscle (Fig. 3*a*). There was no detectable hybridization to brain, liver, or HepG2 RNA. Rat kidney RNA had a weak signal, presumably due to contamination with renal brown fat during RNA purification. An RNA transcript was also present in 3T3-L1 adipocytes but not in pre-adipocytes (Fig. 3*a*). Qualitatively, the tissue distribution of IRGT mRNA was identical to the distribution of the protein as determined by immunoblotting with mAb 1F8 (Fig. 3*c*) or the peptide-specific antiserum, R820 (Fig. 3*d*). In addition, both mAb 1F8 (Fig. 4*b*) and R820 (Fig. 4*a*) immunolabelled a protein of about 45K in 3T3-L1 adipocytes but not in 3T3-L1 fibroblasts. The differences in tissue abundance of the 43K protein immunolabelled with these antisera (brown adipose tissue > heart > red muscle > white adipose tissue and white muscle) parallel the heterogeneity in *in vivo* insulin-stimulated glucose utilization rates that are observed among these tissues^{18–20}. In contrast, using an anti-human erythrocyte glucose transporter antibody, the highest labelling among rat tissues was found in brain (Fig. 3*b*) and a similar concentration of this HepG2-like transporter was detected in both 3T3-L1 fibroblasts and adipocytes (Fig. 4*b*). Relatively weak immunolabelling was observed with the anti-human erythrocyte glucose transporter antibody in muscle, heart and fat. Interestingly, of six heart clones and seven adipocyte clones

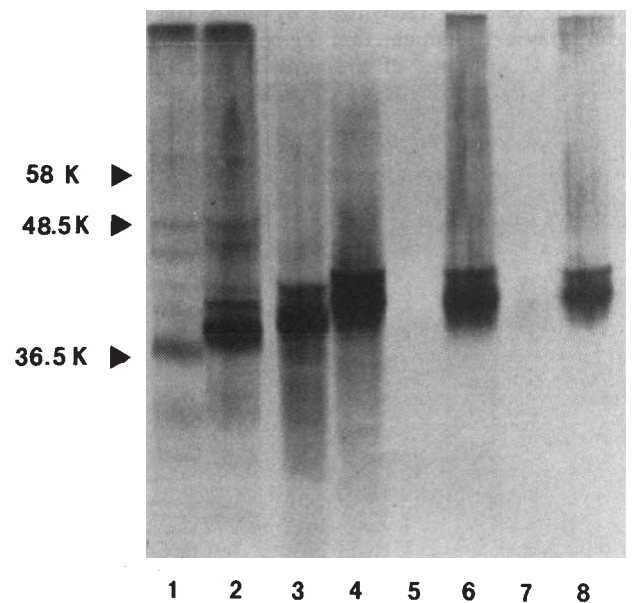


FIG. 2 Translation of IRGT RNA in a reticulocyte cell-free system. A 2,086-bp IRGT cDNA was constructed by ligation of unique *Eco*R1-*Scal* fragments, derived from the two adipocyte clones described in the Fig. 1 legend, at the *Scal* site. The resulting construct was subcloned into the *Eco*R1 site of Bluescribe plasmid. IRGT and HepG2 GT mRNA²⁷ were synthesized *in vitro* using T7 or SP6 RNA polymerase, respectively, and then translated in a reticulocyte lysate system in the presence of ³⁵S-methionine as previously described²⁷. Lysates or pelleted rough microsomes were analysed by SDS-PAGE using an 8% resolving gel. Lane 1, 25 μ l of lysate with HepG2 glucose transporter mRNA; lane 2, 25 μ l of lysate with IRGT mRNA; lane 3, 25 μ l of lysate with HepG2 glucose transporter mRNA translated in the presence of rough microsomes; lane 4, 25 μ l of lysate with IRGT mRNA translated in the presence of rough microsomes. IRGT RNA was translated in the presence of dog pancreatic microsomes, the microsomes were pelleted, washed, solubilized with 1% SDS and diluted (1:10, v:v) with PBS containing 1% Triton X-100 and then divided into two equal aliquots: immunoprecipitation using mAb 1F8 (ref. lane 6), immunoprecipitation using R820 antiserum (lane 8). HepG2 glucose transporter RNA was translated and immunoadsorbed using mAb 1F8 (lane 5) or R820 (lane 7). The relative migration of relative molecular mass standards is indicated on the left.

that were isolated using the HepG2 glucose transporter cDNA, none corresponded to the HepG2-type glucose transporter^{9,10}. Thus, the HepG2-type transporter is probably expressed at low levels, if at all, in insulin-regulatable tissues. The emergence of the IRGT mRNA and protein after differentiation of 3T3-L1 fibroblasts into adipocytes is consistent with the induction of a highly sensitive and responsive hexose transport system under these conditions^{21,22}. These data indicate that 3T3-L1 cells may provide a unique model for studying the factors involved in regulating expression of the insulin-regulatable glucose transporter as well as the cellular factors that confer tissue-specific insulin-dependent glucose transport.

Further evidence that IRGT encodes an insulin-regulatable protein is shown in Fig. 4*a*. The R820 antiserum immunolabelled a protein of similar size to that reported for the adipocyte glucose transporter in the low-density microsomes of untreated rat adipocytes (reviewed in ref. 8). Consistent with previous data obtained using mAb 1F8 (ref. 8), R820 immunolabelled a 43K protein in a plasma membrane fraction from rat adipocytes, the concentration of which was increased about sevenfold upon cellular insulin exposure (Fig. 4*a*). Similar results have been observed with 3T3-L1 adipocytes (data not shown). These data provide further support for the glucose transporter translocation hypothesis^{1,2}. The competition data presented in Fig. 4*c* indicate

that the epitopes recognized by mAb 1F8 and R820 either overlap or are identical, that the epitope of mAb 1F8 is specific for the C terminus of the protein, and that mAb 1F8 and R820 do recognize the same protein in adipocyte membrane fractions.

Several strands of evidence suggest that IRGT encodes the acutely insulin-responsive glucose transporter: (1) the IRGT protein shares considerable sequence and structural identity with other glucose transporters^{9,10,13}, (2) mAb 1F8 immuno-

precipitates the IRGT protein; (3) the tissue distribution of the IRGT protein parallels the acute insulin responsiveness with respect to glucose transport of various tissues *in vivo*; (4) insulin stimulates the translocation of the IRGT protein from a low-density microsomal fraction to the plasma membrane in rat and 3T3-L1 adipocytes; and (5) the differentiation of 3T3-L1 fibroblasts into adipocytes is accompanied by the appearance of IRGT mRNA and protein. Definitive evidence that the cDNA clone described in this study encodes a protein that transports

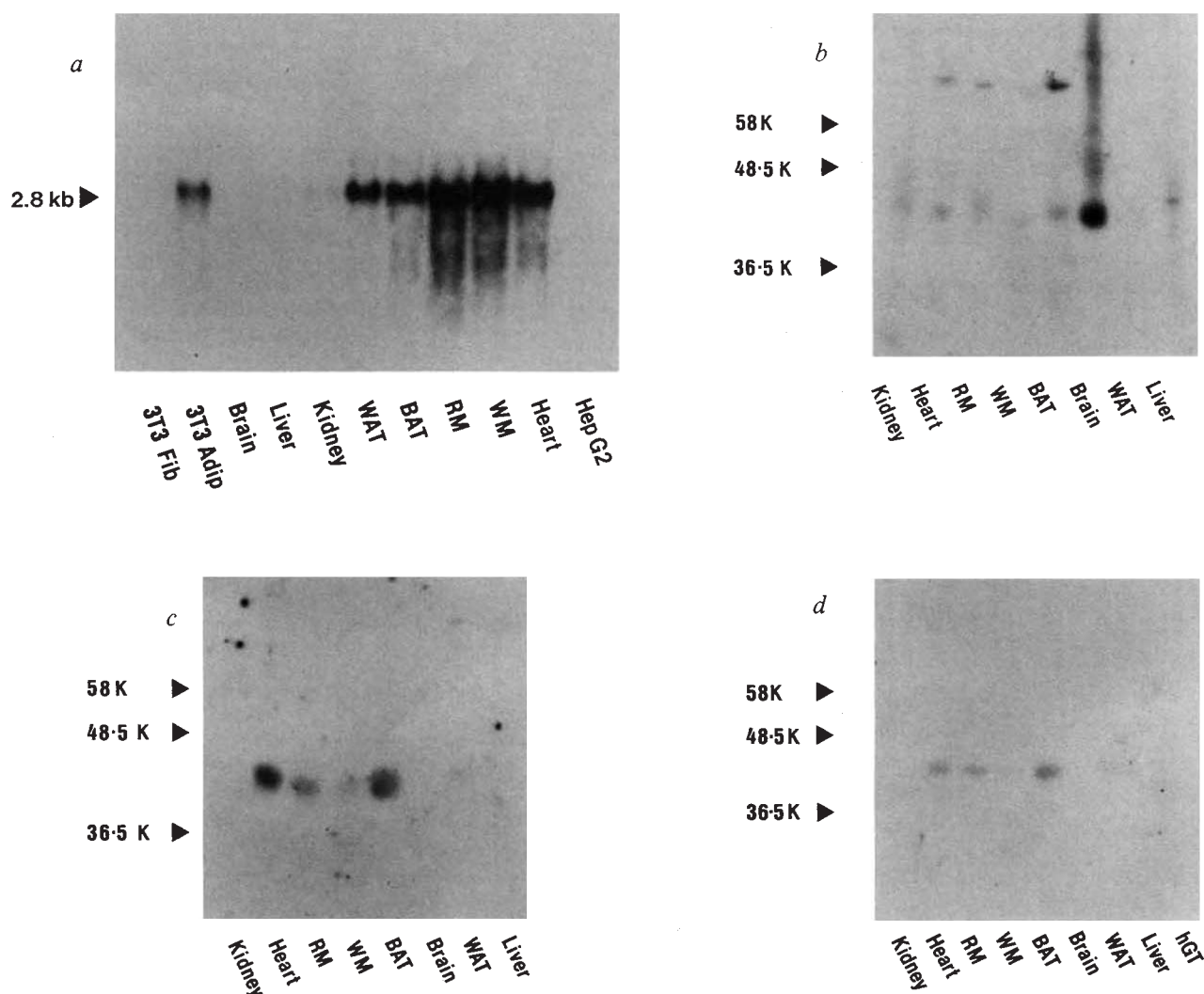


FIG. 3 Tissue distribution of IRGT mRNA and protein. *a*, Total RNA isolated from various rat tissues probed with IRGT cDNA. Abbreviations: 3T3-L1 fibroblasts, Fib; 3T3-L1 adipocytes, Adip; rat white adipose tissue, WAT; rat brown adipose tissue, BAT; rat red skeletal muscle, RM; rat white skeletal muscle, WM. *b*, Rat tissues were homogenized in PBS containing 1% Triton X-100. Total tissue protein (100 µg) from rat tissues were analysed by SDS-PAGE followed by immunoblotting using a polyclonal antisera against the human erythrocyte glucose transporter⁹. Immunoreactive glucose transporter was visualized using ¹²⁵I-labelled Protein A (Amersham). *c*, Total protein extracts (100 µg) from rat tissues were immunoblotted using mAb 1F8. Immunolabelled bands were visualized using ¹²⁵I-labelled sheep anti-mouse IgG (Amersham). *d*, Total protein extracts (100 µg) from rat tissues and 1 µg of purified human erythrocyte glucose transporter (hGT) were immunoblotted using a polyclonal antisera (R820) specific for a peptide derived from the C-terminal sequence of IRGT. Iodo-labelled Protein A was used to visualize the immunolabelled proteins.

METHODS. Total RNA was isolated (*a*) from rat tissues using the guanidinium, CsCl procedure²⁸. The samples (50 µg of each) were electrophoresed on a

formaldehyde-agarose gel (1.2%) and then blotted and fixed onto a nylon membrane. The blot was hybridized to IRGT cDNA labelled with ³²P by nick-translation, washed, and exposed to Kodak XAR-5 film as previously described¹⁹. RNA markers are in an adjacent lane. For *d*, a 12-amino-acid peptide (Thr-Glu-Leu-Glu-Tyr-Leu-Gly-Pro-Asp-Glu-Asn-Asp), corresponding to the 12 terminal amino acids within the C terminus of IRGT (Fig. 1a) was synthesized by the Protein Chemistry laboratory at Washington University School of Medicine. A cysteine was included at the N terminus of the peptide to facilitate coupling of 12.5 mg of the peptide to keyhole limpet haemocyanin (KLH) (40 mg), using the cross-linking reagent m-maleimidobenzoyl-N-hydroxysuccinimide ester (Pierce, Rockford, Illinois), as previously described²⁹. The peptide-KLH conjugate was emulsified in Freund's complete adjuvant and injected at multiple sites intradermally (500 µg) into three New Zealand white rabbits. The rabbits were boosted at three-week intervals with the peptide conjugate (300 µg) emulsified in incomplete adjuvant and bleeds were obtained two weeks after this boost. Sera obtained from all three rabbits tested positive by immunoblotting rat adipocyte subcellular fractions. One rabbit serum (R820) was selected for further study as described here.

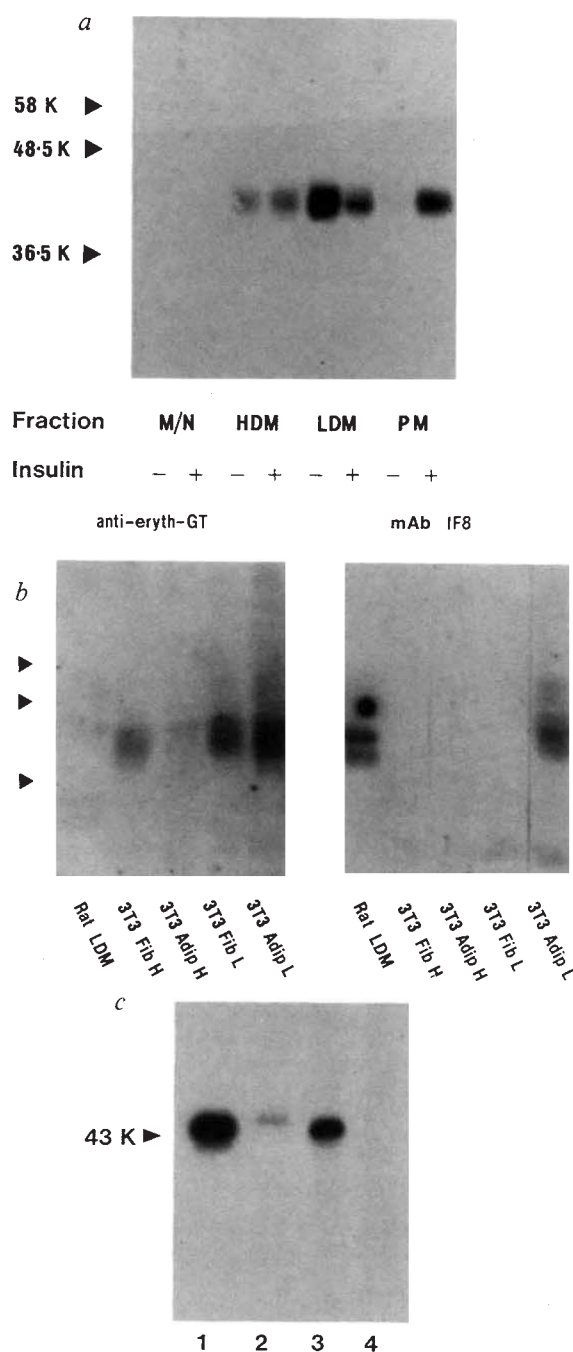


FIG. 4 IRGT encodes an insulin-regulatable protein. *a*, Immunoblot using polyclonal antisera (R820) of subcellular membrane fractions (50 μ g) obtained from rat adipocytes after incubation in the absence or presence of insulin (10^{-8} M) for 15 min: mitochondria/nuclei (M/N), high-density microsomes (HDM), low-density microsomes (LDM) and plasma membranes (PM). *b*, Immunoblot of 3T3-L1 fibroblasts and 3T3-L1 adipocyte membrane fractions (100 μ g) using either a polyclonal antibody specific for the human erythrocyte glucose transporter (anti-eryth GT)⁹ or mAb 1F8. Confluent plates of either 3T3-L1 fibroblasts or adipocytes were homogenized in a buffer containing 20 mM HEPES, 250 mM sucrose, 1 mM EDTA, pH 7.4. Total membrane pellets were layered onto a 1.12 M sucrose cushion, and centrifuged at 100,000g in an SW 41 rotor (Beckman). This procedure resolved membrane fractions above (L) and below (H) the sucrose cushion, of similar protein concentration for fibroblasts and adipocytes. Rat LDM (50 μ g) was electrophoresed in an adjacent lane for comparison. *c*, Rat adipocyte low-density microsomes (20 μ g) were immunoblotted with a polyclonal antisera, R820, at a dilution of 1:500 (lane 1), or with R820 (1:500) in the presence of mAb1F8 (20 μ g ml⁻¹, lane 2), with mAb1F8 alone (2 μ g ml⁻¹, lane 3) or with mAb1F8 (2 μ g ml⁻¹) in the presence of R820 (1:20, lane 4). Immunoreactive species were detected using ¹²⁵I-labelled Protein A (lanes 1 and 2) or ¹²⁵I-labelled sheep anti-mouse IgG (lanes 3 and 4).

glucose will require functional expression. It is interesting that several bacterial proton symporters that do not seem to transport glucose also have significant homology with the HepG2 glucose transporter²³. Thus, it is conceivable that IRGT represents a member of this gene family, whose function involves the transport of a substrate other than glucose. In any case, the overall phenotype of IRGT suggests that this protein has an important role in insulin action. A gene encoding an insulin-regulatable glucose transporter that is expressed only in muscle and fat would be an obvious target for a genetic defect causing predisposition toward insulin-resistant states such as non-insulin-dependent diabetes mellitus. □

Received 11 January; accepted 30 January 1989.

1. Suzuki, K. & Kono, T. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2542–2545 (1980).
2. Cushman, S. W. & Wardzala, L. J. *J. biol. Chem.* **255**, 4758–4762 (1980).
3. Lienhard, G. E., Kim, H. H., Ransome, K. J. & Gorga, J. C. *Biochem. biophys. Res. Commun.* **105**, 1150–1156 (1982).
4. Oka, Y. & Czech, M. P. *J. biol. Chem.* **259**, 8125–8133 (1984).
5. Blok, J., Gibbs, E. M., Lienhard, G. E., Slot, J. W. & Geuze, H. J. *J. Cell Biol.* **106**, 69–76 (1988).
6. Wang, C. *J. biol. Chem.* **262**, 15689–15695 (1987).
7. Oka, Y. *et al. J. biol. Chem.* **263**, 13432–13439 (1988).
8. James, D. E., Brown, R., Navarro, J. & Pilch, P. F. *Nature* **333**, 183–185 (1988).
9. Mueckler, M. *et al. Science* **229**, 941–945 (1985).
10. Birnbaum, M. J., Haspel, H. C. & Rosen, O. M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5784–5788 (1986).
11. Hediger, M. A., Coady, M. J., Ikeda, T. S. & Wright, E. M. *Nature* **330**, 379–381 (1987).
12. Fukumoto, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 5434–5438 (1988).
13. Thorens, B., Sarkar, H. K., Kaback, H. R. & Lodish, H. F. *Cell* **55**, 281–290 (1988).
14. Cairns, M. T. *Biochim. biophys. Acta* **905**, 295–310 (1987).
15. Holman, G. D. & Rees, W. D. *Biochim. biophys. Acta* **897**, 395–405 (1987).
16. Kyte, J. & Doolittle, R. F. *L. Molec. Biol.* **157**, 105–132 (1982).
17. Flier, J., Mueckler, M., McCall, A. & Lodish, H. F. *J. clin. Invest.* **79**, 657–661 (1987).
18. Kraegen, E. W., James, D. E., Jenkins, A. B. & Chisholm, D. J. *Am. J. Physiol.* **248**, E353–E362 (1985).
19. James, D. E., Jenkins, A. B. & Kraegen, E. W. *Am. J. Physiol.* **248**, E567–E574 (1985).
20. James, D. E., Burleigh, K. M., Storlein, L. H., Bennett, S. P. & Kraegen, E. W. *Am. J. Physiol.* **251**, E422–E430 (1986).
21. Rosen, O. M., Smith, C. J., Fung, C. & Rublin, C. S. *J. biol. Chem.* **253**, 7579–7583 (1978).
22. Resh, M. D. *J. biol. Chem.* **257**, 6978–6986 (1982).
23. Malden, M. C. J., Davis, E. O., Baldwin, S. A., Moore, C. M. & Henderson, P. J. *Nature* **325**, 641–643 (1987).
24. Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963–3965 (1983).
25. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1983).
26. Staden, R. *Nucleic Acid Res.* **10**, 4731–4751 (1982).
27. Mueckler, M. & Lodish, H. F. *Cell* **44**, 629–637 (1986).
28. Chirgwin, J. M. *et al. Biochemistry* **18**, 5294–5299 (1979).
29. Lerner, R. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 3403–3407 (1981).

ACKNOWLEDGEMENTS. We thank Drs Lisa Hess, Konrad Keller, John Lawrence, Robert Mercer, Alan Permutt and Karen Tordjman for their helpful suggestions and assistance during this study, Dr John Merlie for providing the rat adipocyte cDNA library, and Kerri James for artwork. Oligonucleotides and peptides were provided by the Washington University School of Medicine Protein Chemistry facility. Support for this work was provided by the Juvenile Diabetes Foundation International, the NIH and the Markey Center for the Molecular Biology of Disease at Washington University Medical School.

Double-strand breaks at an initiation site for meiotic gene conversion

Hong Sun, Douglas Treco, Neil P. Schultes & Jack W. Szostak

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

IT has been proposed that the initiation of meiotic recombination involves either single-strand or double-strand breaks in DNA^{1–3}. It is difficult to distinguish between these on the basis of genetic evidence because they give rise to similar predictions^{4–6}. All models invoke initiation at specific sites to explain polarity, which is a gradient in gene conversion frequency from one end of a gene to the other. In the accompanying paper⁷ we describe the localization of an initiation site for gene conversion to the promoter region of the *ARG4* gene of the yeast *Saccharomyces cerevisiae*. Here, we show that a double-strand break appears at the *ARG4* recombination initiation site at the time of recombination, and that the broken DNA molecules end in long single-stranded tails.

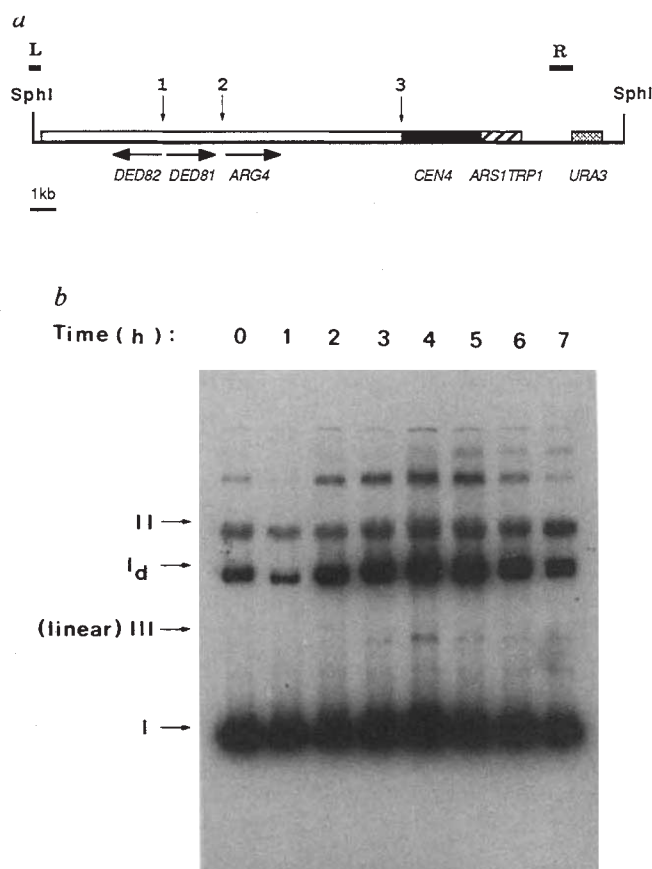


FIG. 1 **a**, Map of plasmid p(SPO13)2 (ref. 27) linearized with *SphI*. The location of the centromere and yeast genes are shown under it. The three meiosis-specific cleavage sites are indicated by arrows. *R* and *L* represent probes used to detect cleaved fragments. The open box represents 15 kb of DNA from the yeast *ARG4* locus. **b**, Time course of undigested plasmid DNA through meiosis. The lowest, most intense band (I) represents supercoiled monomers, band I_d is supercoiled dimer circles, band II is relaxed monomer circles and the transient band III is due to linearized plasmid molecules. The identity of the bands was established by two-dimensional electrophoresis²⁸ and comparison with known markers.

METHODS. Cells of strain NKY278 (*MATa/MATα*, *LYS2::HO/LYS2::HO*, *lys2/lys2*, *ura3/ura3*) carrying plasmid p(SPO13)2 were grown mitotically in YP-acetate medium (1% yeast extract, 2% bacto-peptone, 1% potassium acetate) and transferred to sporulation medium (1% potassium acetate) to start meiosis. Aliquots were removed every hour, frozen, and DNA prepared as described²⁹. The DNA was analysed by electrophoresis on a 0.4% agarose gel, transferred to Nytran membrane and detected with a ³²P-labelled pBR322 (vector-specific) probe.

We have examined the DNA of a yeast plasmid containing a meiotic recombination initiation site for changes related to recombination as the cell progresses through meiosis. The plasmid carries a 15-kilobase (kb) fragment of DNA that includes the *ARG4* gene (Fig. 1a). The diploid yeast strain we used is a derivative of SK-1 (ref. 8), which undergoes a rapid and synchronous meiosis upon transfer to sporulation medium.

Total DNA was extracted from yeast cells collected at hourly intervals after transfer to sporulation medium and the structure of the plasmid DNA was analysed by Southern blotting⁹. Figure 1b shows that the plasmid exists predominantly as supercoiled monomers immediately after transfer to sporulation medium, with some relaxed monomer and various dimer forms. All of these persist through meiosis. As meiosis proceeds however, a new band appears with the mobility of full-length linear plasmid DNA. This band is visible three hours after transfer to sporulation medium, peaks in intensity at four hours, and has decreased

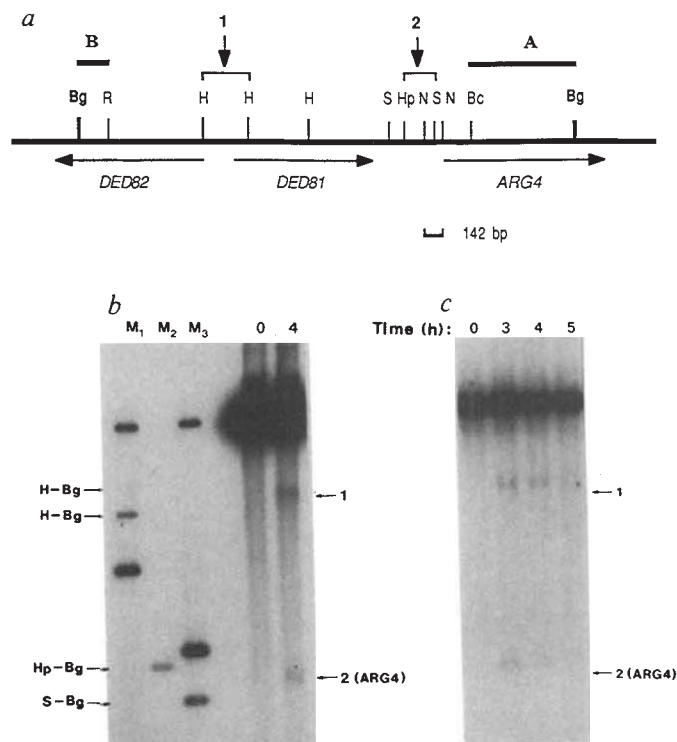


FIG. 2 **a**, Partial restriction map of the *ARG4* region. Restriction sites in and near *ARG4* and two adjacent essential genes, *DED81* and *DED82*, represented by arrows covering their open reading frames. *A* and *B* represent restriction fragments used as probes. *Bc*, *Bc1*, *Bg*, *Bg1*; *R*, *EcoRI*; *H*, *HindIII*; *Hp*, *HpaI*; *N*, *NspHI*; *S*, *SspI*. The 142-bp *NspHI* fragment which was deleted to remove the initiation site is shown below. **b**, Higher resolution mapping of meiosis-specific cleavage sites 1 and 2. DNA from mitotic cells (0 hours), and cells four hours into meiosis was cleaved with *Bgl*II, electrophoresed on a 0.7% agarose gel, blotted and probed with probe *A*. The labelled fragments in the marker lanes extend from the right-hand *Bgl*II site to the restriction sites indicated. Fragments labelled 1 and 2 are seen in the lane containing the 4-h sample but not the 0-h sample lane, and represent cleavage of the *Bgl*II fragment DNA at either site 1 or site 2 respectively. The strongly hybridizing band appearing in both the 0- and 4-h sample lanes represents intact 5-kb *Bgl*II fragment DNA. The *ARG4* cleavage site maps between the *HpaI* and *SspI* restriction sites in the *ARG4* promoter region. The locations of cleavage sites 1 and 2 were confirmed by mapping from the other side of the break using probe *B* (data not shown). **c**, Analysis of chromosomal DNA from strain NKY278 before (0 h) and during (3, 4 and 5 h) meiosis. Cleavages at sites 1 and 2 are still visible, as represented by fragments 1 and 2.

in intensity by five hours. The linear DNA appears at the same time (2.5–5 h) as commitment to meiotic recombination occurs in SK-1 and its derivative strains^{10–11}. In addition, we find that the products of recombination between two restriction site mutations in the *LEU2* gene¹² in an SK-1 strain background are first detectable at four hours, but do not reach their maximal level until five to six hours (data not shown).

We used restriction enzyme digestion to show that the linear plasmid molecules are generated by cleavage at specific sites and to map these sites. Our data indicate that there are three distinct subpopulations of linear plasmid molecules, each resulting from a double-strand break at one of three sites (1, 2 and 3 in Fig. 1a). Each subpopulation consists of about 0.5–1% of the total plasmid population, as determined from comparison with serial dilutions of genomic DNA.

More detailed mapping of the break sites (Fig. 2) revealed that site 2 is located within a 261-base pair (bp) region between the *HpaI* and *SspI* sites immediately 5' to the *ARG4* coding region. Both the *ARG4* transcriptional promoter¹³ and the

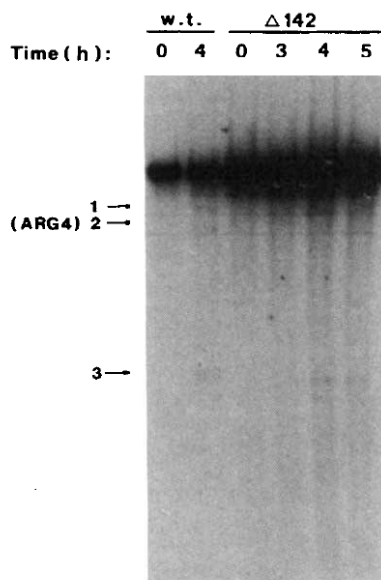


FIG. 3 Effect of deleting the *ARG4* initiation site on the meiosis-specific double-strand break. Plasmid p(SPO13)2- Δ 142 is identical to p(SPO13)2, except that 142-bp are deleted in the *ARG4* promoter region that removes the recombination initiation site⁷. DNA was prepared from cells containing either this plasmid or its wild-type parent (w.t.), before and during meiosis. Samples were digested with *Sph*I, electrophoresed on a 0.5% agarose gel and the fragments detected with probe R (Fig. 1a). The top intense band represents full-length (25 kb) linear plasmid. The three fainter bands are 19, 17 and 10 kb long, and arise from cleavage at sites 1, 2 and 3 respectively. The break in the *ARG4* promoter region (site 2) is absent in the deletion plasmid.

initiation site for meiotic gene conversion (see Fig. 2 in ref. 7) lie in this region. Break site 1 is located within a 450-bp region between two *Hind*III sites, which contains promoters for two divergently transcribed genes, *DED81* and *DED82* (ref. 7; N.P.S. and H.S., unpublished data). Site 3 maps near a functional promoter in the plasmid vector¹⁴.

Deletion of a 142-bp *Nsp*HI fragment in the *ARG4* promoter region decreases gene conversion within *ARG4* from 7.4% to 2% because it removes the *ARG4* initiation site (see ref. 7, Fig. 2, deletion Δ 9). To determine if the same sequence is required for DNA breakage during meiosis, a plasmid derivative carrying the 142-bp deletion was studied in the same strain background. The two plasmids are compared in Fig. 3. The breaks at sites 1 and 3 are unaffected by the deletion and serve as internal controls. The break at site 2 however, is undetectable in the construct carrying the 142-bp deletion.

If the meiosis-specific double-strand breaks are relevant to normal meiotic recombination, they must occur on chromosomal DNA as well as on plasmid DNA. We analysed yeast chromosomal DNA in strain NKY278 in the absence of the plasmid. The breaks at sites 1 and 2 can be detected on chromosomal DNA at the same time interval after transfer to sporulation medium as is observed for plasmid DNA (compare Fig. 2b and c).

We probed the structure of the DNA termini by treating the DNA samples with S1 nuclease, which is specific for single-stranded DNA¹⁵ (Fig. 4). The intensities of the fragments corresponding to breaks at sites 1, 2 and 3 remain constant after S1 nuclease treatment but, remarkably, the size of each fragment decreases. The change in fragment mobility can be seen after 2.5 min of S1 treatment and is complete at 10 min, when all these molecules have been shifted to a faster-migrating species. Prolonged S1 treatment does not result in further degradation of the fragments. The susceptibility of all the broken ends to S1

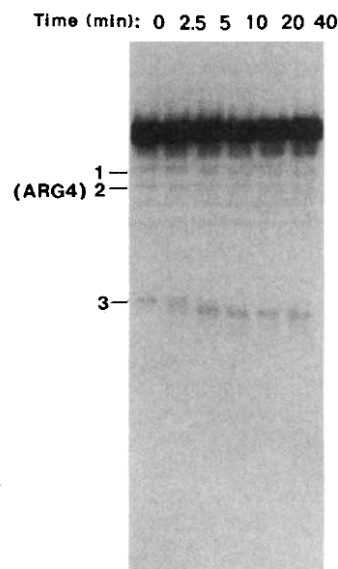


FIG. 4 Treatment of meiotic DNA samples with S1 nuclease. DNA from cells four hours into meiosis was treated with S1 nuclease in buffer containing 200 mM NaCl for the times indicated, digested with *Sph*I, and processed as in Fig. 3. The mobility of the full-length linear plasmid (25 kb) is unaffected by S1, whereas the cleaved fragments 1, 2 and 3 (19, 17 and 10 kb) show an initial rapid decrease in size, indicating that they all end in several hundred bases of single-stranded DNA.

nuclease indicates that all of these molecules terminate with a region of single-stranded DNA. The specificity of S1 nuclease means that this single-stranded region is probably degraded by the nuclease until the enzyme reaches a duplex region¹⁶. When we examined the termini corresponding to the other side of the break, we observed the same S1-sensitive profile (data not shown). Thus, both sides of each of the three breaks end with single-stranded DNA. A similar pattern of sensitivity to S1 nuclease was observed with DNA samples taken at 3, 4, or 5 h into meiosis (data not shown). The single-stranded tails are therefore seen as early as the breaks can be detected. The change in mobility of each of the fragments corresponds to a size difference of a few hundred nucleotides.

To confirm the presence of single-stranded tails on the cut molecules, we made use of the selective binding of single-stranded DNA to nitrocellulose membranes¹⁷. When restricted DNA is transferred directly from the gel without prior alkaline denaturation, fragments representing breaks at sites 1 and 2, are visible on the blot, whereas the intact fragment is absent (data not shown). Control experiments have shown that denatured, but not native, plasmid markers can be detected on these blots. Thus, non-denaturing Southern blots selectively enrich for the population of DNA molecules containing the broken ends, providing strong evidence for these molecules having single-stranded DNA at their termini.

Four arguments can be used to support the involvement of the DNA breaks we observe in the initiation of recombination. First, the DNA breaks are most abundant between three and five hours after transfer to sporulation medium, the time at which meiotic recombination occurs^{10,11}. Second, these breaks occur at specific sites, one of which is the initiation site for meiotic gene conversion at *ARG4*⁷. Third, a deletion that removes the initiation site for meiotic gene conversion at *ARG4* eliminates or greatly diminishes the double-strand break at *ARG4*. Finally, both sides of each of the three breaks end in a few hundred nucleotides of single-stranded DNA. This is consistent with the invasion of duplex DNA by single-stranded DNA being a critical step in the initiation of recombination¹⁻³.

Our data provide strong circumstantial evidence in favour of the double-strand break repair model, but do not eliminate alternative origins for the breaks. First, the double-strand breaks might be formed *in vitro* during DNA preparation. We see the double-strand breaks and their associated single-stranded tails when DNA is prepared by several different procedures, including a very rapid preparation in liquid nitrogen¹⁸. Also, S1 nuclease treatment does not increase the intensity of the fragments, as might be expected if they arose from partial breakage of a single-stranded lesion present *in vivo*. Alternatively, the breaks may arise *in vivo* as consequences of recombination. Holliday junction resolution could lead to duplexes containing nicks on both strands, or mismatch repair could proceed by double-strand gap repair¹⁹. These possibilities are unlikely because we see the breaks in homozygous diploids at the initiation site and as early as three hours into meiosis.

Assuming that the breaks we see are involved in initiation, it is likely that the single-stranded tails are involved in strand invasion, as proposed in the double-strand-break repair model for recombination³. Strand-exchange activities have recently been purified from vegetatively growing and meiotic *S. cerevisiae* cells²⁰⁻²¹. The single-stranded DNA could therefore generate several hundred bases of heteroduplex DNA adjacent to the region of double-strand gap repair. Extensive heteroduplex DNA has been found adjacent to the ends of linear molecules in plasmid transformation systems^{22,23}. The generation of extensive heteroduplex DNA could explain the occurrence of high levels of post-meiotic segregation for certain alleles of *ARG4*^{4,7,24}.

Two of the break sites map within transcriptional promoters, one at the *ARG4* promoter and one between the divergently transcribed genes *DED81* and *DED82*. The third is in the vicinity of a promoter that directs transcription towards the *CEN4* region of the YCp19 vector^{14,25}. Are these promoters identical to the recombination initiation sites? A sequence that stimulates mitotic recombination, *HOT1*²⁶, has recently been shown to be identical to the rDNA promoter. More detailed mapping of the *ARG4* initiation site may resolve the question of the relationship between transcription and recombination initiation. □

Received 2 August 1988; accepted 18 January 1989.

1. Meselson, M. & Radding, C. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 356-361 (1975).
2. Radding, C. M. *et al. Cold Spring Harb. Symp. quant. Biol.* **47**, 821-828 (1982).
3. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. & Stahl, F. W. *Cell* **33**, 25-35 (1983).
4. Fogel, S., Mortimer, R. K. & Lusnak, K. in *The Molecular Biology of the Yeast Saccharomyces* (eds Strathern, J., Jones, E. & Broach, J.) 289-339 (Cold Spring Harbor, New York, 1981).
5. Whitehouse, H. L. K. *Genetic Recombination: Understanding the Mechanisms* (Wiley, New York, 1982).
6. Orr-Weaver, T. L. & Szostak, J. W. *Microbiol. Rev.* **49**, 33-58 (1985).
7. Nicolas, A., Treco, D., Schultes, N. P. & Szostak, J. W. *Nature* **338**, 35-39 (1989).
8. Kane, S. M. & Roth, R. J. *Bact.* **118**, 8-14 (1974).
9. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
10. Resnick, M. A., Chow, T., Nitiss, J. & Game, J. *Cold Spring Harbor Symp. quant. Biol.* **49**, 639-649 (1984).
11. Resnick, M. A., Sugino, A., Nitiss, J. & Chow, T. *Molec. Cell Biol.* **4**, 2811-2817 (1984).
12. Borts, R. H., Lichten, M., Hearn, M., Davidow, L. S. & Haber, J. E. *Cold Spring Harbor Symp. quant. Biol.* **49**, 67-76 (1984).
13. Beacham, I. R., Schweitzer, B. W., Warwick, H. M. & Carbon, J. *Gene* **29**, 271-279 (1984).
14. Marczyński, G. T. & Jeahnig, J. A. *Nucleic Acids Res.* **13**, 8487-8506 (1985).
15. Vogt, V. M. *Eur. J. Biochem.* **33**, 192-200 (1973).
16. Berk, A. J. & Sharp, P. A. *Cell* **12**, 721-732 (1977).
17. Lichten, M. J. & Fox, M. S. *Nucleic Acids Res.* **11**, 3959-3971 (1983).
18. Blin, N. & Stafford, D. W. *Nucleic Acids Res.* **3**, 2303-2308 (1976).
19. Hastings, P. J. *Cold Spring Harbor Symp. quant. Biol.* **49**, 49-53 (1984).
20. Kolodner, R., Evans, D. H. & Morrison, P. T. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5560-5564 (1987).
21. Sugino, A., Nitiss, J. & Resnick, M. A. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3683-3687 (1988).
22. Rothstein, R. *Cold Spring Harbor Symp. quant. Biol.* **49**, 629-637 (1984).
23. Orr-Weaver, T. L., Nicolas, A. & Szostak, J. W. *Molec. cell Biol.* **8**, 5292-5298 (1988).
24. Williamson, M. S., Game, J. C. & Fogel, S. *Genetics* **110**, 609-646 (1985).
25. Stinchcomb, D. T., Mann, C. & Davis, R. W. *J. molec. Biol.* **158**, 157-179 (1982).
26. Voelkel-Meiman, K., Keil, R. L. & Roeder, G. S. *Cell* **48**, 1071-1079 (1987).
27. Wang, H.-T., Frackman, S., Kowalsyn, J., Esposito, R. E. & Elder, R. *Molec. cell Biol.* **7**, 1425-1435 (1987).
28. Wang, J. C., Peck, L. J. & Becherer, K. *Cold Spring Harbor Symp. quant. Biol.* **47**, 85-91 (1982).
29. Borts, R. H., Lichten, M. & Haber, J. E. *Genetics* **113**, 551-567 (1986).

ACKNOWLEDGEMENTS. We thank J. Park, A. Nicolas and A. Ellington for discussion and critical reading of the manuscript, R. E. Esposito for plasmid p(SPO13)2, and W. Raymond, N. Kleckner and M. Lichten for yeast strains. This work was supported by Hoechst A. G.

nature is available in microform.



University Microfilms

International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name _____

Company/Institution _____

Address _____

City _____

State _____ Zip _____

Phone () _____

University
Microfilms
International

Countercurrent chromatography meets MS

Y.-W. Lee and R.D. Voyksner

Interfacing countercurrent chromatography with thermospray mass spectrometry provides a new analytical tool for detecting nonvolatile, hydrophilic or thermally unstable bioactive natural products.

THE development in the 1980s of modern countercurrent chromatography (CCC) instrumentation based upon the fundamental principles of liquid-liquid partition has caused a resurgence of interest in the separation sciences. The advantages of liquid-liquid extraction — the process for separating a multicomponent mixture according to its differential solubility in two immiscible solvents — have long been recognized. In spite of the limitations of the traditional countercurrent distribution methods which prevailed in the 1950s and 1960s, liquid-liquid extraction was used successfully to fractionate commercial insulin into two subfractions differing by only one amide group out of a molecular weight of 6,000 (ref. 1).

In recent years, significant improvements have been made to enhance the performance and efficiency of countercurrent methods². The newly developed high-speed CCC technique utilizes a particular combination of coil orientation and planetary motion to produce a unique hydrodynamic phenomenon in the unilateral phase distribution of two immiscible solvents in a coiled column. The hydrodynamic properties can effectively be applied to perform a variety of countercurrent chromatographies, including true countercurrent³ and foam countercurrent methods⁴. More recently, a 2,000 r.p.m. high-speed CCC has been developed which performs separations with speeds and resolutions comparable to those achieved using HPLC⁵.

We have successfully applied the high-speed CCC system in the separation of plant alkaloids, plant indole hormones, herbicides and bioactive lignins. One of the most distinct advantages offered by CCC is its lack of complications arising from solid supports, such as adsorptive sample loss, deactivation and contamination. Because the two-phase solvent system commonly used in CCC often consists of an organic phase and an aqueous phase, the technique also particularly lends itself to interfacing with mass spectrometry (MS). A solvent with a high aqueous percentage can be used with a volatile buffer such as ammonium acetate to achieve optimal conditions for thermospray MS without suffering the reduction in MS sensitivity typically observed in HPLC/MS with mobile phases low in water content⁶.

MS is generally regarded as the most versatile and universal detection tech-

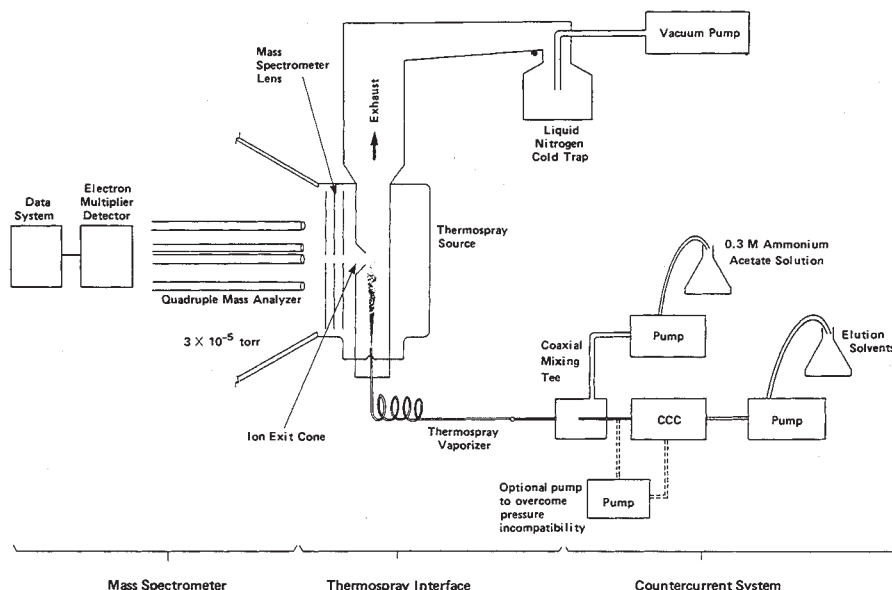


FIG. 1 Schematic diagram of the CCC/MS system.

nique for use with CCC because of its ability to detect compounds with no ultraviolet chromophore and provide precise identification of molecules. The pressure tolerance of high-speed CCC allows it to be interfaced to thermospray⁷ mass spectrometry through an additional pump (to overcome thermospray backpressure). When the resolving power of CCC is coupled in real time with the sensitivity of MS, the individual chromatographic components need not be pure to acquire mass spectra sufficient for identification purposes. Reconstructed single-ion plots in which all of the ions of a given m/z value are plotted against retention time can be used to determine the purity of the total ion current chromatogram (TIC), where the mass spectrometer serves as a compound selective detector. If the intensity of the TIC peak does not rise and fall in accordance with the plot, it is not pure. The on-line coupling of CCC and MS also provides data on trace components which are not easily obtained when the techniques are performed separately.

Configuration details

In our system, shown in Fig. 1, the effluent from a CCC operating at 0.8 ml min^{-1} is introduced into a pump through a zero dead volume tee fitted with a reservoir. Because the CCC produces varying flow rates, the additional pump is operated at 0.7 ml min^{-1} ; the reservoir provides extra solvent or vents excess solvent from the

CCC system as the pressure fluctuates. The effluent from the pump is mixed coaxially with 0.3 M ammonium acetate added at 0.3 ml min^{-1} to create a solvent with sufficient volatility for ion evaporation ionization⁸. The solvent is passed at a rate of 1 ml min^{-1} through a UV detector (254 nm) and into the thermospray interface. (At lower CCC flow rates of $0.3\text{--}0.6 \text{ ml min}^{-1}$ the pressure drop across the thermospray vapourizer is sufficiently low to permit the direct coupling of the CCC effluent to the thermospray interface without the use of the additional pump.)

The thermospray interface connected to the quadrupole mass spectrometer is kept at a source temperature of 250°C and a vapourizer temperature of about 170°C to maximize the chromatography solvent cluster, which has been shown to co-maximize with the analyte under study⁷. No splitting of the CCC effluent is necessary. The solvent is pumped out of the source with a liquid nitrogen cold trap attached to a mechanical rough pump. The analysis is performed employing both negative and positive ion detection using ion evaporation ionization and chemical ionization.

Early results

We have used the thermospray CCC/MS technique successfully in preliminary studies of a mixture of plant alkaloids⁹. Results obtained in our recent application of the technique to the study of bioactive

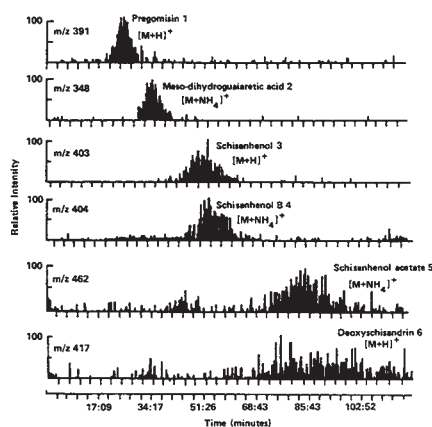


FIG. 2 CCC/MS ion chromatograms of the six major lignins from *Schisandra*. The chemical ionization filament in the quadrupole MS was operated at 1,000 V with a 0.15 mA emission current. The MS was scanned from m/z 180 to m/z 600 in 2 s. The mass calibration of the quadrupole was verified with polypropylene glycol (AMW = 1,000).

lignins¹⁰ and triterpenic acids¹¹ suggest that CCC/MS would be complementary to HPLC/MS as a useful method.

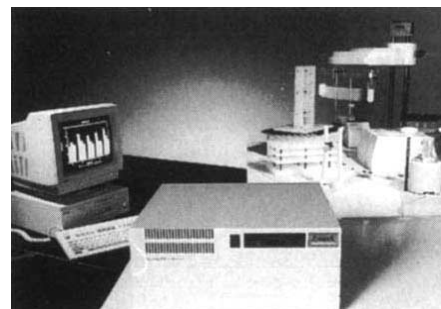
A CCC/UV chromatogram of a crude extract of the bioactive lignin from the herb *Schisandra* shows four distinct peaks. We have used thermospray CCC/MS to confirm the identification of the peaks and further to identify the minor components of the extract. The CCC/MS-selected ion chromatograms (Fig. 2) for the $[M + H]^+$ or $[M + NH_4]^+$ ion of the possible lignins from *Schisandra* indicated six components to be present. The major CCC/UV components present in the extract were confirmed to be schisanhenol and schisanhenol acetate based upon comparison of their thermospray spectra and retention times with the standards for these two components. The minor bioactive lignins such as schisanhenol B4 and deoxyschisanrin 6 were not sufficiently separated from schisanhenol and schisanhenol acetate by CCC with UV detection. But CCC/MS enabled the detection and acquisition of clean spectra of these trace components, which coeluted with the major components in the sample even though they exhibited different characteristic ions. The negative ion spectra for these lignins were 10 to 100 times less intense, and recorded only the major components.

The thermospray CCC/MS total ion current chromatogram presented a different picture of the relative abundances of the bioactive lignins than that obtained using CCC with UV detection. The thermospray CCC/MS response factors of the lignins are more similar than their UV extinction coefficients, enabling a more accurate estimate of the relative percentages of each lignin. Thermospray CCC/MS showed the major active lignins

Pittsburgh preview

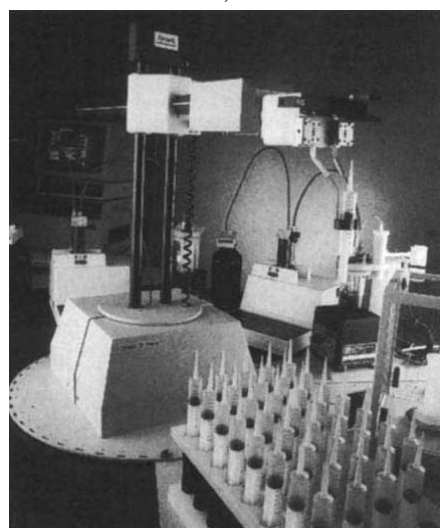
The Pittsburgh Conference will live up to its reputation as one of the largest scientific exhibitions when the doors open in Atlanta, Georgia next week on new products from over 800 companies.

ZYMARK Corporation offers specially configured laboratory robotic systems for automated titrations (Reader Service No. 101). Using Zymark's PyTechnology hardware, the robotic systems can handle the preparation of liquids, viscous liquids and solid samples for Karl Fischer and general titrations. The samples are prepared by the robot, in a sequence that is specified by the operator, and placed directly into the titration cell. The robot starts the titration and monitors all operations to ensure a proper titration. When the titration is done, the results are sent



Zymark's new System V robot controller, for data management and report writing.

back to the robot and a report is generated. Zymark says the robot yields more accurate and precise data than that collected from manual experiments, because its operations are truly reproducible. The company's new System V robot controller performs data management and report writing functions, and archives and transmits data to central computers.



Karl Fischer and general titrations can now be done automatically with Zymark's robot.

from *Schisandra* to be schisanhenol (54 per cent), pregomisin (21 per cent), meso-dihydroguaiaretic acid (13 per cent) and schisanhenol acetate (10 per cent). A total of six lignins were identified: the remaining molecules were nonpolar components with high partition efficiencies in the stationary phase that were not eluted at the same time under the mobile phase conditions.

The coupling of CCC to MS is in its early stages, and new refinements are continually being made. Thick-walled tubing and connections that withstand the high backpressure (roughly 600 p.s.i.) of thermospray MS will make CCC/MS more convenient and easy to use. An analytical CCC system capable of operating at 4,000 r.p.m. is being developed in the Laboratory of Technical Development at the US National Institutes of Health which should raise the separation capabilities of the technique. The feasibility of using two aqueous non-immiscible

International Scientific Instruments is announcing its new computerized scanning electron microscope—the model SX-40A (Reader Service No. 102). The microscope's evacuation system, high-voltage turn-on function and emission settings are all automated for one-touch operation in sequence. The SX-40A's special features include a motor-driven specimen stage, computer-aided auto-focus and auto-stigmator, and automatic contrast/brightness control. The system operates at

solvent systems for separation of proteins is also being investigated. □

Yue-Wei Lee and Robert D. Voyksner are at the Chemistry and Life Sciences Division, and the Analytical and Chemical Sciences Division, Research Triangle Institute, Research Triangle Park, North Carolina 27709, USA. For more information, fill in reader service number 100.

- Hams, J.I., Sanger, F. & Naughton, M.A. *Arch. Biochem. Biophys.* **65**, 427 (1956).
- Ito, Y. *CRC Crit. Rev. analyt. Chem.* **17**, 65 (1986).
- Lee, Y.-W., Cook, C.E. & Ito, Y. *J. Liq. Chromatogr.* **11**, 37 (1988).
- Ito, Y. *J. Liq. Chromatogr.* **8**, 2131 (1985).
- Lee, Y.-W., Ito, Y., Fang, Q.C. & Cook, C.E. *J. Liq. Chromatogr.* **11**, 75 (1988).
- Voyksner, R.D. & Haney, C.A. *Analyt. Chem.* **57**, 991 (1985).
- Blakely, C.R. & Vestal, M.L. *Analyt. Chem.* **55**, 700 (1983).
- Voyksner, R.D., Bussey, J.F. & Pellizzari, E.D. *Analyt. Chem.* **56**, 1507 (1984).
- Lee, Y.-W., Voyksner, R.D., Fang, Q.C., Cook, C.E. & Ito, Y. *J. Liq. Chromatogr.* **11**, 153 (1988).
- Lee, Y.-W. et al. *Analyt. Chem.* (submitted).
- Lee, Y.-W., Fang, Q.C., Ito, Y. & Cook, C.E. 1989 Pittsburgh Conference and Exposition on Analytical Chemistry and Applied Spectroscopy, Abstr. No. 424 (1989).



ISI's computerized SEM is designed for operation at the touch of a single button.

accelerating voltages from 1 kV to 30 kV, allowing it to be used in both beam-sensitive applications and at high beam energies with an ultimate resolution of six nanometers. Samples of up to six inches in diameter can be examined, with scanning speeds variable from 0.2 to 160 seconds. A full selection of options allows the \$45,000 (US) system to be tailored to specific user requirements and applications.

Sartorius Instruments will introduce its new **PLUS performance package for balances** at the Pittsburgh Conference (*Reader Service No. 103*). The instrument has a microprocessor-based keypad with a menu of 16 useful functions that expedite common weighing applications. A few of the functions include the storage of identification numbers, calculation of the



Sartorius's PLUS package gives its range of balances a host of special functions.

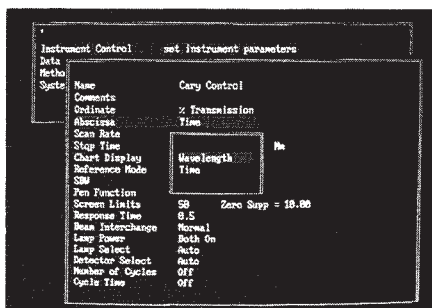
weight of residue in per cent, a fill-toward-zero capability, and even live animal weighing.

Oriel Scientific Ltd has a new catalogue on **lasers and accessories** which contains a valuable composite of new and existing products, with support to build up applied systems (*Reader Service No. 104*). Included are both solid state and conventional HeNe lasers, with full details of supporting optics, launching devices, beam splitters, expanders, mirrors, etalons, optical filters, fibre optics and attenuators. Oriel has pitched the free 80-page catalogue to be useful to both research scientists and applied engineers.

The UV/visible scene

Varian has a new **software interface** program for coupling its Cary 200 series of UV/visible near-IR spectrophotometers

with IBM PC's (*Reader Service No. 105*). The \$1,250 (US) PC Control program provides centralized control of the spectrophotometer from an IBM PC/AT or PS/2 model 30 or 50. Key features of the package include the ability to store and retrieve instrument parameters, display data in real time, produce multicolor pen plots, and convert data to standard formats usable by other IBM-compatible



A screen display from Varian's software interface for its UV/visible near-IR spectrophotometers.

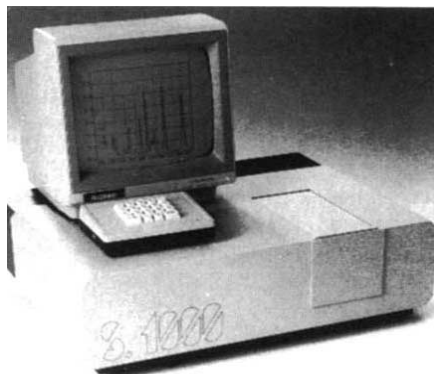
software packages. PC Control has a graphical interface which allows the user to scroll through display options, confirm selections, and move step-by-step through the software.

Perkin-Elmer's **Lambda 6 PC-controlled UV/visible spectrophotometer** was designed for applications in environmental, pharmaceutical and biochemical analyses (*Reader Service No. 106*). The Lambda 6 PC system is comprised of the double-beam, scanning Lambda UV/visible spectrophotometer integrated with an industrial-standard personal computer and software. Perkin-Elmer says the optical design of the system ensures low stray light, spectral fidelity, fast scan speeds, baseline stability and low noise. The system's software automatically processes and stores data from procedures such as scanning, time drive, wavelength programming, spectral processing and kinetics studies. Its open-ended design allows future expansion to increase data handling capabilities.



Perkin-Elmer's Lambda 6 UV/visible spectrophotometer is controlled by a computer.

Secomam is introducing a **computer-driven spectrophotometer** at the Pittsburgh Conference, designed to scan the entire UV/visible range at speeds of up to 2,400 nm per minute (*Reader Service No. 107*). In addition to standard measurement modes, Secomam's model S.1000 can perform kinetic measurements at a single wavelength with sampling rates of as high as 30 per second. The system's software can carry out sophisticated calculations — including derivatives up to the fourth order — curve storage, and curve expansion. Up to five components of a single solution can be quantitatively analysed automatically with the system. A multi-standard calibration function permits four different curve fitting models and as many as twelve different combinations of math-



Secomam's S.1000 spectrophotometer also comes with a microprocessor instrument.

ematical axes. The cell compartment can accommodate cells with light paths up to 100mm long, for highly sensitive detection of low level components.

New **quantitative analysis and enzyme kinetics software** is now available from Hewlett-Packard for its model HP 8452A diode array UV/visible spectrophotometer and ChemStation (*Reader Service No. 108*). The HP 89511A quantitation software package simplifies single component method development and multicomponent analyses. In single component analyses, the software allows the selection of

wavelengths which provide the best calibration curve and the greatest reproducibility. The concentrations of up to twelve components can be determined within a known sample matrix. The software allows the use of absorption and derivative spectra, and the company says it saves time in method development by mathematically synthesizing spectra before measuring actual samples to determine the feasibility of performing the multicomponent analysis. The new enzyme kinetics package, model HP89512A, provides biochemists with the means to investigate enzyme reactions, rate data and reaction models. Nine mathematical models are available, or users can develop models from their own equations. The kinetics of several reactions may be measured in parallel with the help of a new multi-cell transport mechanism for the HP spectrophotometer.

Capillary electrophoresis

Bio-Rad is now offering an instrument for performing **high performance capillary electrophoresis** (Reader Service No. 109). The new model HPE 100 system can perform high-resolution separation and analysis of a variety of samples, including peptides, proteins, DNA fragments, amino acids, bacteria and viruses. The



Bio-Rad's benchtop unit for performing capillary electrophoresis.

system is a compact benchtop unit which contains an electrophoresis chamber and all of the electronics in a safe interlocked cabinet. Separation takes place in a short, narrow-bore capillary tube with a patented coating which Bio-Rad says eliminates electroendosmosis. Samples are injected with a microliter syringe and the electropherogram results are produced on either a strip chart recorder, an integrator, or a PC-driven data collector within about ten minutes, according to the company. Electrophoresis methods that can be adapted to HPE include free zone, displacement and discontinuous methods, and isoelectric focusing.

Applied Biosystems is also introducing a new system for **high-pressure capillary electrophoresis** — its new model 230A



Applied Biosystems' entry onto the capillary electrophoresis market.

HPEC system (Reader Service No. 110). The capillary electrophoresis process enables the user to separate, visualize, quantitate and recover mixtures of biological macromolecules. The 230A HPEC is a micro-preparative instrument for the separation and purification of proteins, DNA, and synthetic oligonucleotides. Applied Biosystems says the system offers the resolving power of gel electrophoresis with the convenience of continuous sample elution, real-time visualization and direct sample collection. In addition, HPEC collects peaks of interest for further analysis or purification. As samples are eluted from the tube gel, a continuous flow of elution buffer carries the separated sample components through an on-line detector and into an integral fraction collector. The temperature-controlled gel compartment accepts a variety of column diameters and lengths, and the programmable power supply offers multiple operation modes. The company says standard electrophoresis protocols may be adapted readily to the tube gel format of the system.

Chromatography currents

Sanki Laboratories has a new model LLB benchtop, high-speed **analytical centrifugal partition chromatograph** (Reader Service No. 111). Centrifugal partition chromatography (CPC) is a new liquid chromatographic technique which utilizes liquid-liquid partition and counter-current distribution to fractionate complex mixtures of chemical substances. Because CPC is not based on a solid stationary support, there is no pH limitation, irreversible retention or catalytic molecular



Sanki's newest analytical centrifugal partition chromatograph runs at 2,000 r.p.m.

rearrangement of labile species. The model LLB's main centrifugal module can be operated at speeds of up to 2,000 r.p.m. The unit contains 8,000 partition channels for the high-efficiency analysis of mixtures ranging from natural products and pharmaceuticals, to antibodies and genetically engineered products.

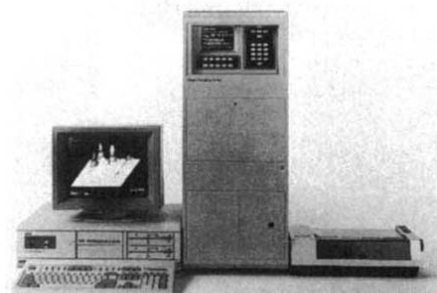
Dionex has announced what may be the first system for **environmental analysis** to bring together the power of HPLC, ion chromatography and flow injection analysis (Reader Service No. 112). The basic system consists of a quaternary gradient module, a chromatography module and a postcolumn module. A column-switching valve allows a single sample to be analysed using various combinations of variable wavelength detectors, fluorescence detectors, a conductivity detector and a pulsed amperometric detector. Post-column derivatization adds the ability to determine contaminant amounts down to p.p.b. levels. The system can be quickly adapted for flow injection analysis by replacing the chromatographic column with a method-specific AppliCard. The fluid path in all of the system components is metal-free, so all types of solvents can be used without corrosion concerns.

A new **capillary SFC method development kit** is now available from Lee Scientific (Reader Service No. 113). The kit contains four 2.5-m, 50 μ m i.d. columns containing one of each of the following four phases: SB-Octyl-50, SB-Biphenyl-30, SB-Cyanopropyl-50 and SB-Smectic-50. Lee says the 2.5-m columns produce quick results during screening runs, yet deliver efficiencies high enough for the resolution of complex samples, so that they may be rapidly screened to determine which stationary phase is best suited for a particular application.

Perkin-Elmer has a new **liquid chromatography system** which offers computer control for methods development and automated sample processing (Reader Service No. 114). The Perkin-Elmer LC Analyst system includes a quaternary LC pump, random-access autosampler, and a personal computer with LC Analyst software. It will accept data from any detector, and is designed to assist chromatographers in the development of methods even when little or no sample information is available. Perkin-Elmer says the LC Analyst's ability to process large numbers of samples automatically and thorough documentation allow developed methods to be transferred easily to routine use.

Special features of the LC Analyst system include customized methods development, software packages, history file reports for the documentation of methods and real-time changes, and automatic systems startup and shutdown.

The Waters chromatography division of Millipore has designed a Peptide Analyzer for high resolution **peptide mapping** (Reader Service No. 115). The HPLC system features dual flow paths optimized

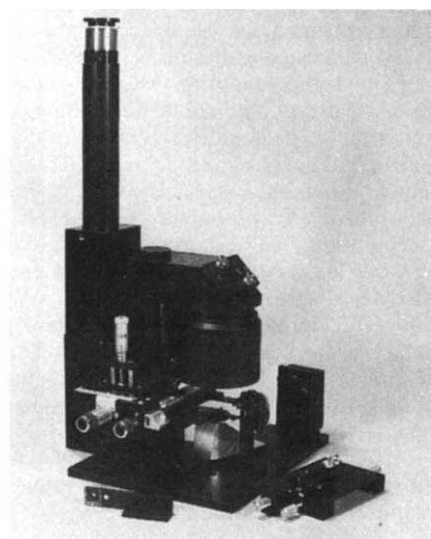


Waters' system for use in peptide mapping.

for microbore and analytical scale operation. The analyzer comes with the Power-Line HPLC controller, which allows single-point control of gradient, injection and detection parameters. The system is also equipped with the Waters model 990+ Photodiode Array Detector which offers complete UV/visible spectral analysis. Spectra can be directly compared to those acquired previously using the detector spectral library. Waters says the system is ideal for the preparation of picomole amounts of pure peptides for sequence determination, peptide maps for the quality control of recombinant DNA products, or for qualitative structural analysis.

Analytical spectroscopy

Spectra Tech has redesigned its FT-IR Spectra-Scope to allow the **microsample analysis** of crystals, fibres, laminates, con-



Spectra Tech's 'scope has been redesigned for microsample analysis.

taminants, and other small samples difficult to load onto a standard beam condenser (Reader Service No. 116). The scope has a $15\times$ collecting objective and a $10\times$ ocular lens which together magnify the sample area up to 150 times. The company says the Spectra-Scope has a low signal-to-

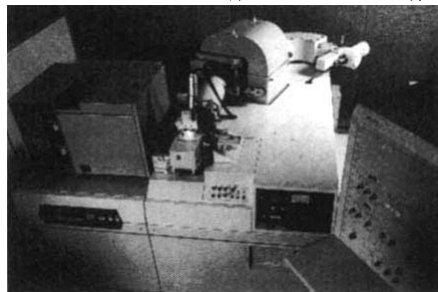
noise ratio, even with a sample as small as $15\ \mu\text{m}$. Samples are positioned using the X-Y-Z translation stage and light straying is minimized with a remote aperture which masks the image of the sample. Illumination is provided by a high intensity lamp with intensity adjustment. The new \$8,500 (US) Spectra-Scope is designed for use with most FT-IR spectrometers.

Applied Research Laboratories is launching a new version of its thermally stabilized SpectraSpan VII **multi-element spectrometer**, available in either simultaneous or sequential configuration (Reader Service No. 117). In simultaneous mode, ARL says the SpectraSpan VII is capable of determining as many as 20 elements in 30 seconds; sequentially, it can determine more than 70. The computerized, bench-top spectrometer is controlled by an IBM PS/2 computer and MS-DOS software that provides operator feedback, generates internal reports, and documents all analytical diagnostic information. The spectrometer features automatic dynamic background correction and high resolution, and a new digital LCD display for gas flow control and profiling functions. The SpectraSpan VII can be used in environmental, clinical, agricultural and geological laboratories.

Philips is introducing its model PW1480 **sequential X-ray spectrometer** — a versatile system for applications requiring high performance and flexibility in the analysis of solids, powders and liquids (Reader Service No. 118). Developed from the PW1400 series, a new model incorporates a number of proven features: a high precision goniometer, a 60 kV or 100 kV generator, and a wide range of single- and dual-anode X-ray tubes. With the model PW1480, Philips has added a wider choice of collimator and crystal options, plus a bidirectional, eight-position crystal changer to facilitate custom tailoring. To go with the spectrometer, Philips offers its new X-40

analytical software package, which has facilities for both quantitative and qualitative spectrometer operation, and a user interface which supports both the command and menu modes. To make the software easier to use, Philips has included logical defaults at input points throughout the program and easily accessible "help" information. Philips says the software package reduces calculation time by allowing automatic qualitative and semiquantitative analysis of several samples with a single command.

Jeol has a new **mass spectrometer with ion optics** which it says is as equally suited for high-sensitivity, high-resolution GC/MS as it is for high-molecular weight

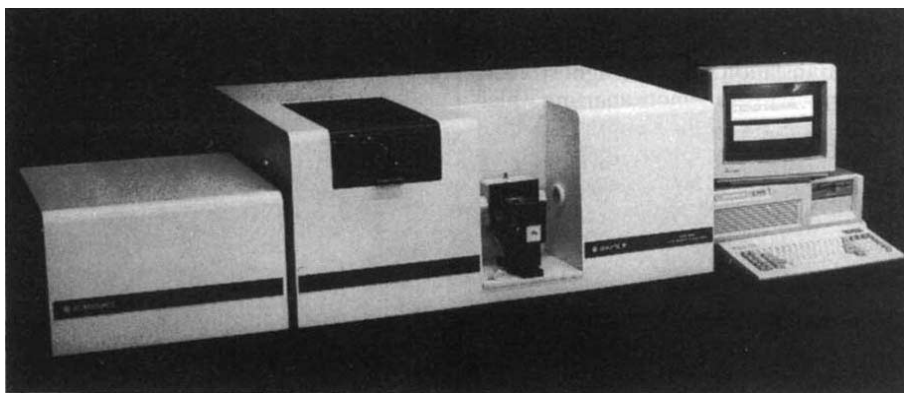


Ion optics of Japanese design enhance Jeol's new mass spectrometer, the model Sx-102.

LC/MS measurements (Reader Service No. 119). The Jeol SX-102 can be used to detect very small traces of drugs, pesticides and pollutants with the SIM GC/MS technique. Its LC/MS applications using the frit inlet system include the analysis of amino acids, peptides, proteins and pharmaceuticals. The instrument's QHQOC optics geometry, created by Matsuda of the University of Osaka in Japan, provides high ion transmission and positively shapes the beam to allow a slit width five times that of conventional mass spectrometers, for high sensitivities at both low and high resolution. The instrument has a practical mass range of 3,000 daltons at 8 kV, but Jeol says it can measure compounds of greater than 5,000 daltons, such as bovine insulin.



Philips' sequential X-ray spectrometer can handle solids, powders and liquids.

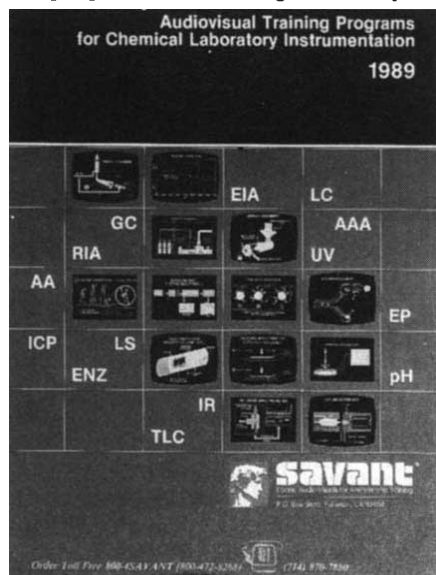


Analyte's answer for multielement atomic absorption spectrophotometry.

Analyte Corporation is introducing its new model SL-16 **multielement atomic absorption spectrophotometer**, which the company says is capable of determining in rapid sequence the concentrations of up to 20 elements in either liquid or solid samples (*Reader Service No. 120*). The SL-16 contains a fast-moving turret which holds 16 hollow cathode lamps combined with a monochromator programmed to move to the correct wavelength while the turret is rotating. Analyte claims the system can determine 20 elements from one sample in under two minutes. Conventional burner systems can be used with the spectrophotometer for liquid samples; Analyte's Atomsource glow-discharge device is required to atomize solid samples into the argon gas stream. The Atomsource can be connected to the spectrophotometer in roughly 10 minutes. The SL-16 is controlled by an AT-type computer that can be programmed with up to 400 different analytical methods, each of which includes lists of elements, wavelengths, and slit widths.

Information resources

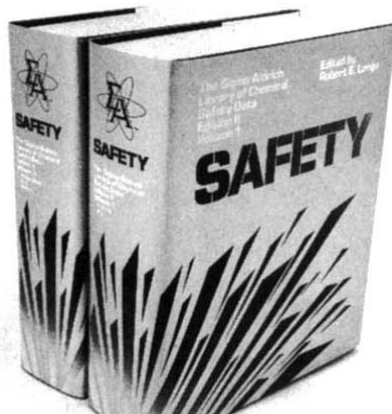
Savant's 1989 catalogue of audiovisuals for **chemical laboratory training** contains 76 programmes covering a variety of



Savant's catalogue on tutorials.

analytical techniques for both industrial and medical applications (*Reader Service No. 121*). Available training programmes range from peptide analysis and UV/visible spectroscopy, to general laboratory procedures and chemical safety. Most Savant programmes are available as both slide and cassette tape shows and videotapes in Beta or VHS formats. The programmes are professionally narrated, and vary in length from 20 to 45 minutes.

The Sigma-Aldrich Company has published the second edition of its two-volume library on **chemical safety** (*Reader Service*



Sigma-Aldrich's volumes contain things chemists need to know for a safer lab.

No. 122). Over 14,000 chemicals and 24,000 products are listed in the two books, with information ranging from chemical structure and physical properties to health hazards and recommended waste disposal methods. The volumes together contain 4,098 pages, and cost \$325 (US).

Finnigan MAT is offering a free **colour wallchart** to be used in the interpretation of peptide and protein mass spectra (*Reader Service No. 123*). The information provided on the wallchart includes amino acid structures, molecular weights, and hydrophobicity indexes. The wallchart also lists the gas phase fragmentation pathways of peptide chains and the differences between their nominal, average and exact molecular weights. The colour wallchart is the first of a series

planned by Finnigan MAT to explain the biochemical applications of mass spectrometry; the company will introduce additional wallcharts later this year.

To complement its "How To" notes detailing applications and techniques for using FTIR spectrometers, Mattson



Mattson's FT-IR spectroscopy videotape.

Instruments Ltd is releasing an **education film** entitled "Infra-Red Spectroscopy" (*Reader Service No. 124*). The video-cassette film is produced in conjunction with the British Broadcasting Company's Open University programme, and offers both the established and the potential user theoretical and practical information on the FTIR spectroscopy technique. A range of methods for the measurement of gas, solid and liquid samples are demonstrated on the video, using Mattson FTIR spectrometers.

The 1989 **catalogue** is now available on the Eyela line of scientific instruments from Tokyo Rikakikai Company Ltd. (*Reader Service No. 125*). The 45-page catalogue lists a host of products for analytical chemistry, cell culture, microbiology, and biotechnology. Besides such basic laboratory items as freeze-dryers, water baths, incubators, ovens and water purifiers, the Eyela catalogue spells out Tokyo Rikakikai's specialty systems for laboratory- or production-scale rotary evaporation, cell immobilization and carboxylic acid analysis. The catalogue also describes the company's line of batch and airlift bioreactors for plant and animal cells, and fermentation pilot plants. For chromatography needs, the catalogue includes information on the Eyela droplet counter-current chromatograph, flash chromatograph, rotation locular counter-current chromatograph and liquid-liquid counter-current extractor.

Beckman is introducing its new Data Logger **software for spectrophotometry** at the Pittsburgh Conference, and is exhibiting the package along with one of its DU Series 60 spectrophotometers (*Reader Service No. 126*). Data Logger allows users to log discrete absorbance readings of up to 114 samples on 3.5-inch diskettes, collect wavelength scans, carry data to



Beckman's new Data Logger software expands the capabilities of its spectrophotometers.

another IBM PC or PS/2 personal computer for analysis, and convert data to Lotus format. With the software, users can add or subtract spectra, and display up to six spectra on the screen for direct viewing and comparison. The Data Logger package includes an IBM PS/2 model 25 computer with 640 kBytes RAM and dual 3.5-inch disk drives, DOS version 3.3, the Beckman DU Data Capture software program, a formatted data diskette and necessary cables. Data Logger also supports the Beckman DU-50 and DU-70 series spectrophotometers, and the DU-6 and DU-7 UV-visible spectrophotometers.

Biology at Pittcon

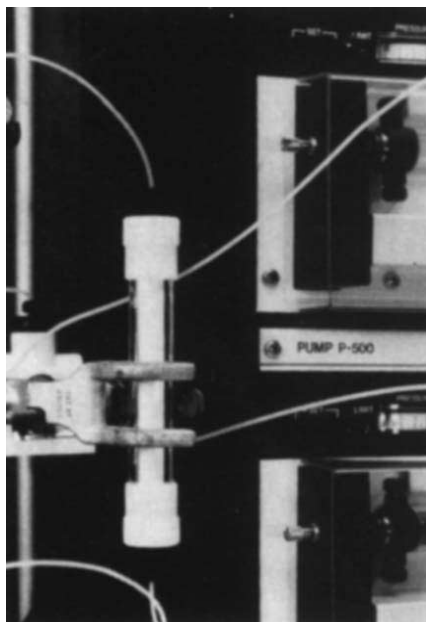
OROS Systems, Inc. is launching its second instrument for **protein purification** — the touch-screen operated MultiLab system (*Reader Service No. 127*). OROS designed the MultiLab for use in small-scale preparative purifications and for determining optimal purification methods. The operator chooses the preferred method by touching the MultiLab's touch-sensitive screen, and the an expert software system takes over to monitor the purification and control all parameters. The \$35,000 (US) MultiLab runs at both low and medium pressures, and includes air sensors to detect the absence of buffers and an air eliminator to protect the column. OROS Systems says the system can be used with immunopurification, affinity chromatography, hydrophobic interaction chromatography, ion-exchange, gel-filtration or chromatofocusing techniques.

The Generator **DNA synthesizer** will be on display in DuPont's booth at the Pittcon show (*Reader Service No. 128*). The \$22,500 (US) Generator uses β -cyanoethyl-diisopropyl phosphoramidites, which DuPont says demonstrate consistently high base-to-base coupling efficiencies. The precision fluid meter of the Generator delivers microlitre quantities of reagent to the reaction column to save reagents, and can perform at least 100 coupling cycles before requiring replenishment, according to the company. The Generator has a built-in system check which alerts the user and places itself on hold if either fluid flow or gas pressure fall outside set parameters. Operating the system requires no computer language

skills: nucleotide sequences are entered directly from the keyboard, and the screen display shows the status of the reaction cycle and the base position in the sequence. A custom programming function allows users to develop their own chemistry programs, which can be stored on disk through the system's floppy disk drive. DuPont offers a complete range of pre-measured reagents for use with the Generator which come in thumb-screw-on bottles that cannot be misinstalled.

Polymer Laboratories Ltd have a new line of **LC columns** for protein analyses which link the company's biocompatible metal-free column with a range of quaternized PL-SAX strong anion exchange materials (*Reader Service No. 129*). The PL-SAX media are composed of a quaternary amine applied to a rigid macroporous polymer, giving a pH range of 1–13 and a high degree of physical stability.

Molecular Devices recently introduced the first laboratory instrument to employ biosensor technology — the Threshold **total DNA assay** system based on the light-addressable potentiometric silicon sensor developed by Hafeman, Parce and McConnell (*Reader Service No. 130*). The Threshold system quantitates contaminating DNA in samples of biotechnology products, down to the picogram level. In the first step, the sample is incubated for one hour with a single reagent that contains two binding protein conjugates: a single-stranded binding protein linked to a hapten specific for the capture membrane, and an anti-DNA monoclonal antibody joined to urease. The sample is then transferred to the vacuum unit of the Threshold, which concentrates it onto eight measurement sites on the Threshold biosensor



Polymer Lab's all-glass setup for the chromatography of biomolecules.

stick. When the stick is dipped into the reader, it comes in contact with urea, and the ensuing enzyme reaction changes the local pH at each measurement site, which alters the surface potential of the sensor proportionally to the amount of DNA present. Molecular Devices says all DNA fragments larger than 100 bases can be quantitated with the computer-driven Threshold system, which can process up to 32 samples in roughly 15 minutes. The company is currently developing other DNA probe and immunoassay kits for use with the Threshold.

Japanese imports

Asahi Chemical sells a **desk-top desalinator** called the Micro Acilyzer based on electrodialysis with an ion-exchange membrane (*Reader Service No. 131*). The electrodialysis technique employs a pair of cation- and anion-exchange membranes which are activated by an electric potential to transport negatively and positively charged ions out of the sample. Asahi says its Micro Acilyzer can be used for the desalination of oligosaccharides, and the removal of phosphate buffer from peptides and nucleosides. The company sells two versions of the Micro Acilyzer: model G1 processes 10 ml samples, and model G3 can desalinate up to 1,000 ml.

The model SDM5500 **dynamic mechanical spectrometer** from Seiko Instruments can be used for the measurement of the physical properties of solid materials (*Reader Service No. 132*). The instrument consists of a rheometer, an analysis module and an X-Y plotter. Stress from the force generator is applied by a probe, and the strain caused in the sample is measured as the displacement signal by a strain meter, and sent to the computer for processing. The system's software incorporates windows so that running conditions and data can be viewed simultaneously. □

ADVERTISEMENTS

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID

Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5–10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

POB 126, D-8000 Munich 65, FRG

Reader Service No.45

Move the work or move the people?

Richard Pearson

Pressures on the crowded and expensive south-east corner of England are leading to 'satellite offices' and long-distance commuting. Many companies see relocation as the only answer.

DESPITE a fast-growing economy, great regional disparities remain in the United Kingdom. In many of the inner cities and towns in the North, unemployment remains stubbornly high, often reaching levels in excess of 15 per cent. Yet in many parts of the South-East, registered unemployment has fallen to under 5 per cent and skill shortages are growing worse. In an effort to alleviate the worst effects of these shortages employers are now looking to recruit from farther afield, in some cases extending their catchment areas by a few miles, in others encouraging long-distance commuting from several hundred miles away. Others are setting up remote networking centres while others are relocating in order to find the staff they need.

A recent study of the South-East has shown the different stages of distance recruitment clearly¹. The sizes of the individual local catchment areas vary greatly, increasing with the seniority of the appointment, the nearness to the centre of London, and to a lesser extent transport links. The better the links the greater the distances people are prepared to travel.

Perks?

The first response of employers to shortages is to consider applications from beyond the usual spatial area. This usually involves advertising in the local media, and is a fast, easy, low-cost option. Companies are often reluctant to make even this simple move, fearing poor time keeping and increased absenteeism as a result of long-distance travel. In London, 64 per cent of the professional staff surveyed spend more than an hour travelling to work, compared to 7 per cent in the rest of the South-East and even fewer in some other parts of the country (see table). Towns such as Bristol, Coventry, Doncaster and Ipswich, are now joining the South-Coast resorts such as Brighton and Eastbourne as regular commuter towns for London. Surprisingly, outside of the construction industry, few firms offer improved financial or other assistance with commuting costs such as season ticket loans or car purchase schemes. These are still seen as perks rather than as recruitment aids.

The next step for recruiters is rather bigger and involves advertising in more distant labour markets. Here, a major problem is choosing which labour market might be appropriate, the main indicators used being the prevailing level of local unemployment, local employment struc-

ture and reports of redundancies in firms similar to their own. Firms with branch networks, however, are able to use them as information sources and recruitment agents, and also encourage staff transfers into the South-East. High-tech firms tend to use consultants.

House prices

There remains, however, a growing reluctance on the part of individuals to move into London and the South-East, with regional house price differentials being a major barrier². In contrast with their reluctance to improve the assistance for longer distance commuting, most firms surveyed are improving their relocation package, although only a minority are investing in sheltered housing, and assisting in property disposal and search, property leasing schemes and joint equity schemes. The numbers of staff benefiting from the latter are small. For most companies the response is financial, providing reimbursement for the costs involved in the move as well as for higher costs in the new area, and supplements to compensate for the disruption. Few companies have detailed knowledge of the real costs involved, but their estimates of expenditure were in the range £5,000-£10,000 with a tax exemption

opportunities. About half of them were employed on temporary or short-term contracts, few were allowed to adjust their working patterns to fit in with their travel arrangements. The extra costs of this travelling tend to fall on the individual, with little subsidy or inducement coming from the employers³.

Another variation on long-distance working that is now being pioneered in the United Kingdom is that of network information technology centres in inner cities, remotely serving customers in the South-East. Recruitment to these centres is likely to be from local training schemes and it is intended that the centres will not only be commercially viable, but will also provide valuable work experience and further training to those who would be otherwise employed, and boost the skills profile of the local community. The technology is already proven, the key is to attract the customers. In the United States a number of companies now send all their low-level data processing work via satellites to centres in the Caribbean where wage rates are way below those of the United States — in some cases below even the cost of office rental in North America.

An increasing number of employers, including central government, are also

Travel-to-work times in selected regions

Journey time	London	South-East	North	Wales
0-29 min	7%	62%	57%	72%
30-59 min	29%	30%	36%	26%
60+ min	64%	7%	7%	2%

Data from ref. 1

limit of £8,000 last year. (Tax exemption is now being increased to £17,200.)

Although many of the available jobs are in the South-East, financial, housing and social reasons often mitigate against a move to the region. Now there is a growing class of professional long-distance commuters, with as many as 10,000 people regularly travelling from the North to work in London and the South-East. Interviews with a small sample of these commuters show that many are middle managers continuing in jobs or careers that they had had earlier in their lives, but that they are now earning far higher salaries than they were previously and demand high levels of job satisfaction. A major motivation for many 'long-distance commuters' interviewed was a history of redundancy, not that of long-term unemployment, and the lack of local job

moving the whole workplace and setting up offices outside the South-East to capitalize on their property assets, at the same time alleviating skill shortages. With labour supply and working conditions unlikely to improve in the South East in the foreseeable future, relocation of the work, whether in terms of the whole office or just part of the workload via telecoms links, looks like a more effective long-term solution than trying to attract more people to travel or move to the work. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Falmer, Brighton BN1 9RF, UK.

1. *Relocation and Recruitment Difficulties of Employers in the South East* (Institute of Manpower Studies, 1988).
2. *Nature* **332**, 98 (1988).
3. *Britain's New Industrial Gypsies* (Policy Studies Institute, 1989).